

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____, Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 07-15-02

Searcher: Beverly E 4997

Terminal time: 20

Elapsed time: _____

CPU time: _____

Total time: 25

Number of Searches: _____

Number of Databases: 1

Search Site

_____ STIC

_____ CM-1

_____ Pre-S

Type of Search

_____ N.A. Sequence

_____ A.A. Sequence

_____ Structure

_____ Bibliographic

Vendors

_____ IG

_____ STN

_____ Dialog

_____ APS

_____ Geninfo

_____ SDC

_____ DARC/Questel

_____ Other CGN

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:16:23 ; Search time 58.93 Seconds
(without alignments)
1053.627 Million cell updates/sec

Title: US-09-298-924-6
Perfect score: 2978
Sequence: 1 MTFAYKIDGNEVIFTWAPY.....POHIEGKYFDKGFALYKL 559

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	2978	100.0	559	17 AAR90619	Sulfolobus solfata
2	1748.5	58.7	556	17 AAR92755	Trehalose-releasin
3	1748.5	58.7	556	17 AAR90620	Trehalose-releasin
4	1089.5	36.6	595	22 AAG92072	Sulfolobus acidoca
5	1089.5	36.6	610	22 AAR80206	Corynebacterium gl
6	1060.5	35.6	598	16 AAR77471	Trehalose releasin
7	1058	35.5	597	16 AAR80290	Trehalose releasin
8	1025	34.4	589	16 AAR80289	Trehalose releasin
9	1025	34.4	596	16 AAR77470	Trehalose releasin
10	1024.5	34.4	575	21 AAY85157	Trehalose-releasin
11	1024.5	34.4	575	21 AAY85165	Trehalose-releasin

12	399	13.4	652	16	AAR80037	Bacillus stearothe
13	399	13.4	652	17	AAR96109	Starch-branching-e
14	385.5	12.9	648	10	AAP94635	B. thuringiensis a
15	374.5	12.6	639	12	AAR11271	B. stearothermophil
16	371.5	12.5	718	22	AAB60903	Propionibacterium
17	361	12.1	630	22	AAB69074	Aquifex aeolicus v
18	348	11.7	670	22	AAU60913	Propionibacterium
19	348	11.7	766	20	AAU00870	S. tuberosum isoam
20	346.5	11.6	658	13	AAR23787	Heat-resistant pul
21	345.5	11.6	666	20	AAV34991	C. pneumoniae prot
22	342.5	11.5	793	20	AAU00869	S. tuberosum isoam
23	340	11.4	731	22	AAG91100	C glutamicum prote
24	340	11.4	731	22	AAB79423	Corynebacterium gl
25	339	11.4	764	21	AAV50819	Wheat isoamylase p
26	334	11.2	606	18	AAW36602	S. tuberosum debra
27	333	11.2	1250	21	AAV91279	Group B Streptococ
28	331.5	11.1	618	20	AAV37184	Protein involved i
29	331.5	11.1	798	20	AAW73552	Pullulanase protel
30	331.5	11.1	1938	20	AAW73553	Full length Pullul
31	329.5	11.1	818	20	AAW17523	Zea mays SU1 starc
32	329.5	11.1	893	17	AAW09257	Bacillus alkaline
33	329.5	11.1	1938	17	AAW09255	Bacillus alkaline
34	315	10.6	772	19	AAW49871	Thermotoga maritim
35	312.5	10.5	621	21	AAW19285	A polypeptide with
36	312	10.5	772	18	AAW34567	Thermotoga maritim
37	309.5	10.4	722	20	AAV35095	C. pneumoniae prot
38	303.5	10.2	829	22	AAE05691	Bacillus acidopull
39	303.5	10.2	862	21	AAV78513	Pullulanase (pulB)
40	303	10.2	931	20	AAV27357	Group B Streptococ
41	300	10.1	498	10	AAV96105	Sequence of any C
42	299	10.0	928	22	AAE05719	Bacillus deramific
43	296	9.9	762	21	AAW90977	N. denitrificans a
44	295	9.9	555	14	AAW43340	Alpha-glucosidase
45	294.5	9.9	785	19	AAW70886	Protein encoded by

ALIGNMENTS

RESULT 1
ID AAR90619 standard; Protein; 559 AA.
XX AAR90619;
AC AC
DT 29-JUN-1996 (first entry)
XX
DE Sulfolobus solfataricus amylase for alpha, alpha-trehalose prodn.
XX
KW transferase; amylase; Sulfolobus; production; alpha, alpha-trehalose;
KW malto-oligosaccharide; hydrolysis.
XX
OS Sulfolobus solfataricus.
XX
PN W09534642-A.
XX
PD 21-DEC-1995.
XX
PF 14-JUN-1995; 95WO-JP01189.
XX
PR 21-APR-1995; 95JP-0120673.
PR 15-JUN-1994; 94JP-0133354.
PR 18-AUG-1994; 94JP-0194223.
PR 31-OCT-1994; 94JP-0290394.
PR 21-NOV-1994; 94JP-0286917.
PR 21-NOV-1994; 94JP-0311185.
XX
PA (KIRI) KIRIN BEER KK.
XX
DR WPI; 1996-049671/05.
DR N-PSDB; AAT12325.
XX
PT Sulfolobus spp. derived transferase and amylase - for production of

PT alpha, alpha-trehalose from malto-oligosaccharide(s)

PS Claim 108; Page 235-240; 357pp; Japanese.

CC The amylase is derived from Sulfolobus solfataricus. The amylase acts on a saccharide having at least three sugar units, which are pref. glucose units at the reducing end (the linkage between the first and second glucose units is alpha-1, alpha-1, while the linkage between the second and third glucose units is alpha-1,4), to hydrolyse alpha-1,4 linkages within the sugar chain, yielding alpha, alpha-trehalose and also mono- and disaccharide hydrolysis products. The amylase has a mol. wt. of 61 to 64 kDa. It is characterised by working at pH 4.5-5.5 and at 60-85 deg.C. It has an isoelectric point of 4.3-5.4 and retains at least 100 percent activity after 6 hrs. at 80 deg.C. It is completely inhibited by 5 mM copper sulphate. Use of a transferase and the amylase in succession on suitable substrates such as malto-oligosaccharides, is useful for the production of alpha, alpha-trehalose.

XX Sequence 559 AA;

Query Match 100.0%; Score 2978; DB 17; Length 559;
Best Local Similarity 100.0%; Pred. No. 3.6e-236;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTFAYKIDGNEVIFTLWAPYQKSVKLVLEKGLYMERDEKGYFTITLNNKVRDRYKYV 60

Db 1 mtfaykidgneviftlwapyqskvlkgleglymerdekgyftitlnnvkvrdrkyv 60

Qy 61 LDDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKEDLIIEIHVGFTPEGTFF 120

Db 61 lddaseipdpasryqpegvhgpsqligeskefnnetflkediieihvgftpegtffe 120

Qy 121 GVIRKLDYLKDLGITAIEIMPTAOPGKRWDGYGVLYAVQNSVGGPEGRKLVDEAHK 180

Db 121 gvirkldylkdlgitaieimptaqpgkrwdgygvlyavqnsyggpegrfkvldeahk 180

Qy 181 KGLGVILDVYVNHVGPENYVVKLGPFYSQYKTPWGLTFNFDDESDVSRKFIENVEY 240

Db 181 kglgvildvvyvnhvgpegnymvklgpyfsqyktpwgltnfddaesdevrkfienvvey 240

Qy 241 WIKENYVDGFRDLDAVHAIDTSPKHILEIADVVKYNRIVIAESDLNDPRVNPKECG 300

Db 241 wikenyvdfgrldavhaidsphilleeadvvhkynriviaesdlndprvnpkecg 300

Qy 301 YNIDAQWDDFHHSHAYLTGERQGYTDFGNLDDIVKSYKDVYDGKYNFRKRTHGE 360

Db 301 ynidaqwddfhshihayltgerqgytcdgnlddivksykdvydgkynfrkrthge 360

Qy 361 PVGELDGCNFVYVYIIONHDQVGNRKGRIIKLVDRSYKIAAALYLLSPYIPMIFMGEY 420

Db 361 pvgelgcnfvyvylqnhdqvgngkgeriiklvdrsykiaaalyllspyipmifmgey 420

Qy 421 GEENPFYFFSDFSKLLTQVREGKKGQDTPDQDSTFNASKLSWKIDEEIFSFKI 480

Db 421 geenpfyffsdfskllqvgregkkengqtdpddstfnasklskwkideeifsyki 480

Qy 481 LTKRKELSIACDRRVNVNGENWLLIKGREYFSLYVFSKSIIEVKYSGTLLSSNNSFP 540

Db 481 likmrkelsiacdrvrnvngenvwllikgreyslyvfskssievkysgtllssnsnfp 540

Qy 541 QHIEEGKYEFKGFALYKL 559

Db 541 qhieegkyefkdgfalykl 559

RESULT 2

AAR92755
ID AAR92755 standard; Protein; 556 AA.

XX AAR92755;

AC AAR92755;

XX DT 03-AUG-1996 (first entry)

XX Trehalose-releasing thermostable enzyme.

DE Thermostable enzyme; trehalose; sweetener.

KW Sulfolobus acidocaldarius strain ATCC 33909.

XX OS A09527131-A.

XX 01-FEB-1996.

XX 21-JUL-1995; 95AU-0027131.

XX 04-JUL-1995; 95JP-0189760.

XX 21-JUL-1994; 94JP-0190180.

XX 11-APR-1995; 95JP-0109128.

XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

XX Kubota M, Mitsuizumi H, Sugimoto T;

XX WPI: 1996-106284/12.

XX N-PSDB; AAT16899.

XX Recombinant thermostable enzyme from Sulfolobus acidocaldarius, -

XX releases trehalose from non-reducing saccharide at temps. exceeding

XX 55 degrees Centigrade

XX Claim 2; Page 53-54; 74pp; English.

XX A thermostable enzyme (AAR92755) of Sulfolobus acidocaldarius ATCC

XX 33909 releases trehalose from non-reducing saccharides having a

XX trehalose structure as an end unit and a degree of polymerisation of

XX at least 3. It has a mol.wt. of 54-64 kDa (SDS-PAGE), a pI of

XX 5.6-6.6 and is substantially not inactivated when incubated in aq.

XX solution (pH 7.0) at 85 deg for 60 min. Recombinant enzyme is

XX obtd. by expression of an isolated DNA fragment (AAT16899) in

XX host cells, pref. Escherichia coli, using e.g. vector paluescript

XX II SK(+). The trehalose is useful as a sweetener.

XX Sequence 556 AA;

XX Query Match 58.7%; Score 1748.5; DB 17; Length 556;

XX Best Local Similarity 58.9%; Pred. No. 4.1e-135;

XX Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

Qy 2 TFAYKIDGNEVIFTLWAPYQKSVKLVLEKGLYMERDEKGYFTITLNNKVRDRYKYV 61

Db 3 sfgnieknkgiflwapyvnsvklk-lskklipmekndegfveiddeenitysyil 61

Qy 62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNETFLKEDLIIEIHVGFTPEGTFF 121

Db 62 edkreipdpasryqpglvhdkslirtdyqildgkvkiedliiylhvgftsgqgnfkg 121

Qy 122 VIRKLDYLKDLGITAIEIMPTAOPGKRWDGYGVLYAVQNSVGGPEGRKLVDEAHK 181

Db 122 vieklldylkdlgitaieimptaqpgkrwdgygvlyavqntyyggpewelaklvneahkr 181

Qy 182 GLGVILDVYVNHVGPENYVVKLGPFYSQYKTPWGLTFNFDDESDVSRKFIENVEY 241

Db 182 giavildvvyvnhigpegnylglgpyfsdryktpwgltnfddrgcdqvrkfilenvvey 241

Qy 242 IKEYNVDDGFRDLDAVHAIDTSPKHILEIADVVKYNRIVIAESDLNDPRVNPKECG 301

Db 242 fktfkidgfrldavhaidsphilleeadvvhkynriviaesdlndprkiv--kddcgy 299

Qy 302 NIDAQWDDFHHSHAYLTGERQGYTDFGNLDDIVKSYKDVYDGKYNFRKRTHGE 361

Db 300 kidaqwddfhshihayltgerqgytcdgnlddivksykdvydgkynfrkrthge 359

Qy 362 VGEIDGCNFVYVYIIONHDQVGNRKGRIIKLVDRSYKIAAALYLLSPYIPMIFMGEY 421

Db 362 vgeidgcnfvyvylqnhdqvgngkgeriiklvdrsykiaaalyllspyipmifmgey 421

Db 360 vgdpprkfvvqhndqvgngngerslsltdkttlylmaatlylspylplfmgeey 419

Qy 422 EENPFYFSDSKLIQGVREGKKGQDTPQDDESTFNASKLSWKIDIEIFSFKIL 481

Db 420 etnpfffsdpvlikvreglrkennqmidpqqseafllksklskideevldyykql 479

Qy 482 IKMKELSIACDRRVNVNGENWLIIGREYFSLYVFSKSIEVYKSGTLLSSNNSFPQ 541

Db 480 inirkryn-nckrvkevrrregncitlmeikigiliasfddivinskitgnlligi--gfpk 536

Qy 542 HIEBKG-YEFDKGFALYKL 559

Db 537 klkdelikvnrvgvyql 555

RESULT 3

AAR90620

ID AAR90620 standard; Protein; 556 AA.

AC AAR90620;

XX 29-JUN-1996 (first entry)

XX Sulfolobus acidocaldarius amylase for alpha, alpha-trehalose prodn.

XX transferase; amylase; Sulfolobus; production; alpha, alpha-trehalose;

KW malto-oligosaccharide; hydrolysis.

XX Sulfolobus acidocaldarius.

XX W09534642-A.

XX 21-DEC-1995.

XX 14-JUN-1995; 95WO-JP011189.

XX 21-APR-1995; 95JP-0120673.

PR 15-JUN-1994; 94JP-0133354.

PR 18-AUG-1994; 94JP-0194223.

PR 31-OCT-1994; 94JP-0290394.

PR 21-NOV-1994; 94JP-0286917.

PR 21-NOV-1994; 94JP-0311185.

XX (KIRI) KIRIN BEER KK.

PA WPI; 1996-049671/05.

XX N-PSDB; AAT12326.

XX Sulfolobus spp. derived transferase and amylase - for production of

PT alpha, alpha-trehalose from malto-oligosaccharide(s)

XX Claim 112; Page 244-250; 357pp; Japanese.

XX The amylase is derived from Sulfolobus acidocaldarius. The amylase acts

CC on a saccharide having at least three sugar units, which are pref.

CC glucose units at the reducing end (the linkage between the first and

CC second glucose units is alpha-1, alpha-1, while the linkage between the

CC second and third glucose units is alpha-1,4), to hydrolyse alpha-1,4

CC linkages within the sugar chain, yielding alpha, alpha-trehalose and also

CC mono- and disaccharide hydrolysis products. The amylase has a mol. wt. of

CC 61 to 64 kDa. It is characterised by working at pH 4.5-5.5 and at 60-85

CC deg.C. It has an isoelectric point of 4.3-5.4 and retains at least

CC 100 percent activity after 6 hrs. at 80 deg.C. It is completely inhibited

CC by 5 mM copper sulphate. Use of a transferase and the amylase in

CC succession on suitable substrates such a malto-oligosaccharides, is

CC useful for the production of alpha, alpha-trehalose.

XX Sequence 556 AA;

Query Match 58.7%; Score 1748.5; DB 17; Length 556;

Best Local Similarity 58.9%; Pred. No. 4.1e-135;

Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

Qy 2 TFAYKIDGNEVIFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVVRDRYKYL 61

Db 3 sfngnieknkgfklwapyvnsvklk-lskklpmekndegffeideiddeenltsyil 61

Qy 62 DDASEIPDASRYOPEGVHGSPQIIQESKEFNFTFLKEDLIIEYHVGFTPEGTFEG 121

Db 62 edkreipdasryqplgvhdksqlirtdyqildlgkiedliiyeihvgftsfqgnfk 121

Qy 122 VIRKLDVLDGIFTAIEIMPTAQPPGRKMGYDGVLYAVQNSYGGPGRKLVDEAHKK 181

Db 122 vieklldvldgltgietmpaqfpgordwgydgvfilyavqntyggpbwelaklvneahr 181

Qy 182 GLGVILDVYVNHVPEGNMYKLGPFYSQRYKTPWGLTFNFDDESVRKRKFILENEYW 241

Db 182 giavilgvvynhigpegnylllgpyfsdryktpwgtfnfdrgcdqvrkfilenveyw 241

Qy 242 IKEYNVDFRLDAVHAIIIDTSPKILBEIADVHKYRIVIAESDLNDPRVVPNKEKCY 301

Db 242 fktfkidglrladvhaifdnspkhlqelaeakahqlgkfviaesdlndpkiv--kddcgy 299

Qy 302 NIDAAQWDDFHHSTHAYLTGEROGYVTDGNLDDIVKSYKDVYVYDCKYSNFRKTHGEP 361

Db 300 kidaqwddthhahafitkekdydyqdfriediektkdvfydgkysrygrthgap 359

Qy 362 VGLDGCNFVYVYQNDQVGNRKGRIIRIKLVRESYKIAAALYLLSPYIPMIPMGEEY 421

Db 360 vgdipprkfvyfqnhdqvgngngersltdkttlylmaatlylspylplfmgeey 419

Qy 422 EENPFYFSDSKLIQGVREGKKGQDTPQDDESTFNASKLSWKIDIEIFSFKIL 481

Db 420 etnpfffsdpvlikvreglrkennqmidpqqseafllksklskideevldyykql 479

Qy 482 IKMKELSIACDRRVNVNGENWLIIGREYFSLYVFSKSIEVYKSGTLLSSNNSFPQ 541

Db 480 inirkryn-nckrvkevrrregncitlmeikigiliasfddivinskitgnlligi--gfpk 536

Qy 542 HIEBKG-YEFDKGFALYKL 559

Db 537 klkdelikvnrvgvyql 555

RESULT 4

AAG92072

ID AAG92072 standard; Protein; 595 AA.

AC AAG92072;

XX 26-SEP-2001 (first entry)

XX C glutamicum protein fragment SEQ ID NO: 5826.

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis.

XX Corynebacterium glutamicum.

XX EP1108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-0127688.

XX 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

Db 175 vheaaygpeaygrfvdaaaglgvqdvynhlgpsgnylprfgpylkqgegnwtgds 234
Qy 221 NFDDAEDVAKFLENYEYKIKENYKGVDFRLDAVHAIDTSPKHILKEIADVVHKNRI 280
Db 235 nldpgsdhvrriyldnlamwldryvdglrlidavhallderavhiledfgaladqisae 294
Qy 281 V-----IAESDLNDPRVVPNPKKGYNIDQWDDFHHSIHAYLTGERQGYTDFGNLD 334
Db 295 vgrplilaesdlndprllyprdvngyglgqwsddfhavhvnvtgettygsdfdsala 354
Qy 335 DIVSKYKDVYDGYKSYNFRKTHGCEPVGELDGCNF-----VVYQNDHQVGNRGK 386
Db 355 alakvlrdgfhdsysfrhrhgrpi-----nfavhpaalvcsqnhdqgnratg 408
Qy 387 ERITKLDVRESYKTAALYLLSPYPMIFMGEEYGEENPFYFSDSKLIQGVREGK 446
Db 409 drlseqtpygsalaaavltlcpftpmllmgeeygastpwwfftshepelgkataegri 468
Qy 447 KE---NGOD-----TDPODESTFNASKLSWKIDEE-----IFSFKILIKMRKELSIAC-- 492
Db 469 kefermgwdpavvpqdpqpetfrskldwaaeaegdharlielyrsltalrrstpdltkl 528
Qy 493 ---DRRVNVNGENWLIIGREYFSLYVFSKSSIEVKYSGTLLL 533
Db 529 gfdetqafdedarwlrfrgvgvllnfseqpvsldgagtail 572

RESULT 7

AAR80290
ID AAR80290 standard; Protein: 597 AA.
XX
AC AAR80290;
XX
DT 19-JAN-1996 (first entry)
XX
DE Trehalose releasing enzyme.
XX
KW Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW alpha-maltotriosyltrehalose; alpha-maltotetraosyltrehalose;
KW maltopentaosyltrehalose; sweetener; taste-improving agent;
KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW pharmaceuticals.
XX
OS Arthrobacter sp. Q36.
XX
PN EP671470-A2.
XX
PD 13-SEP-1995.
XX
PF 07-MAR-1995; 95EP-0301474.
XX
PR 07-MAR-1994; 94JP-0059840.
PR 07-MAR-1994; 94JP-0059834.
XX
PA (HAYB) HAYASHIDARA SEIBUTSU KAGAKU.
XX
PI Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX
DR WPT; 1995-312772/41.
DR N-PSDB; AAQ98670.
XX
XX DNA encoding a trehalose releasing enzyme - which releases trehalose
XX from a non-reducing saccharide having a trehalose structure as an
XX end unit.
XX
PS Claim 3; Page 24-25; 45pp; English.
XX
CC This enzyme can be used for the preparation of trehalose with high
CC yields and efficiency from non-reducing saccharides such as
CC alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC alpha-maltotriosyltrehalose, alpha-maltotetraosyltrehalose and
CC maltopentaosyltrehalose. The trehalose can be used as a sweetener,
•

CC taste-improving agent, quality-improving agent, stabiliser, filler,
XX excipient or adjuvant in food products cosmetics and pharmaceuticals.
SQ Sequence 597 AA;

Query Match 35.5%; Score 1058; DB 16; Length 597;
Best Local Similarity 40.5%; Pred. No. 2.5e-78;
Matches 229; Conservative 88; Mismatches 191; Indels 58; Gaps 15;

Qy 14 FTLWAPYQSKVKLVLEKG-LYEMER-----DEKGVFTI-----TLNNKVRDYRKYKVLVD 62
Db 18 ydwapnaesvtil--laggeryamrraetgpdagwtaagaptgdndv----ygyld 71
Qy 63 -DASEIDPPASRYOPEGVHGSPSQIIOESK-EFNNETFLKKE--DLIIYEIHVGFTPEGT 118
Db 72 gdepldprrtrpdpghalsrtfdpsayswqddawgrelgavayelhlgtftpegt 131
Qy 119 FEGVIRKLDYLKDLGITAIEIMP IAQFPGRKRWGVDGYLYAVQNSYGGPEGFRKLKLVDEA 178
Db 132 leaaagkllylaglgvdfiellpvnafngthwgydgwfvavhedyggpeayqrfvdaa 191
Qy 179 HKKGLGVILDVVYVNHVGPENYMKLGPYFSQKYKTPWGLTFNEDDAESDEVRKFIENLV 238
Db 192 haaglgvldvvyvnhlgpsgnylprfgpylkqgegnwtgdsvnldgpgsdhvrriyldnl 251
Qy 239 EYWIKEYNVGDFRLDAVHAIDTSPKHILEIADVVHKNRIV-----IAESDLNDPRV 292
Db 252 amwldryvdglrlidavhallderavhiledfgaladqisaevgrplilaesdlndprl 311
Qy 293 VNPKEKCGYNIDAQWDDFHHSIHAYLTGERQGYTDFGNLDDIVKSKYKDVYVYDGYKSN 352
Db 312 lypcdvngyglgqwsddfhavhvnvtgettygsdfdsalaalakvlrdgfhdsyys 371
Qy 353 FRKTHGCEPVGELDGCNF-----VVYQNDHQVGNRGKGERIILKLVDESKYKIAAAL 404
Db 372 frerhgrpi-----nfavhpaalvcsqnhdqgnratgdrisltlpygsalaaavl 425
Qy 405 YLLSPYPMIFMGEEYGEENPFYFSDSKLIQGVREGKKE---NGOD-----TDPOD 457
Db 426 tltpftpmllmgeeygastpwwfftshepelgkataegrikefermgwdpavvpqdp 485
Qy 458 ESTNASKLSWKIDEE-----IFSFKILIKMRKELSIAC-----DRRVNVNGENWLI 507
Db 486 petfrskldwaaeaegdharlielyrsltalrrstpdltklgfdetqafdedarwlr 545
Qy 508 KGREYFSLYVFSKSSIEVKYSGTLLL 533
Db 546 irggvqvllnfseqpvsldgagtail 571

RESULT 8

AAR80289
ID AAR80289 standard; Protein: 589 AA.
XX
AC AAR80289;
XX
DT 19-JAN-1996 (first entry)
XX
DE Trehalose releasing enzyme.
XX
KW Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW alpha-maltotriosyltrehalose; alpha-maltotetraosyltrehalose;
KW maltopentaosyltrehalose; sweetener; taste-improving agent;
KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW pharmaceuticals.
XX
OS Rhizobium sp. M11.
XX
PN EP671470-A2.
XX
PD 13-SEP-1995.
XX

```
PF 07-MAR-1995; 95EP-0301474.
XX
PR 07-MAR-1994; 94JP-0059840.
PR 07-MAR-1994; 94JP-0059834.
PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX
XX WPI: 1995-312772/41.
XX N-PSDB; AAQ98669.
XX
XX DNA encoding a trehalose releasing enzyme - which releases trehalose
XX from a non-reducing saccharide having a trehalose structure as an
XX end unit.
XX
XX Claim 3; Page 21-22; 45pp; English.
XX
XX This enzyme can be used for the preparation of trehalose with high
XX yields and efficiency from non-reducing saccharides such as
XX alpha-glucosyltrehalose, alpha-maltosyltrehalose,
XX alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
XX maltopentaosyltrehalose. The trehalose can be used as a sweetener,
XX taste-improving agent, quality-improving agent, stabiliser, filler,
XX excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
XX Sequence 589 AA;
XX
XX
XX Query Match 34.4%; Score 1025; DB 16; Length 589;
XX Best Local Similarity 40.3%; Pred. NO. 1.3e-75;
XX Matches 229; Conservative 80; Mismatches 197; Indels 62; Gaps 16;
XX
XX QY 14 FTLWAPYQKSVKLVLEKG-LYEMER-----DEKGYFTITLNNVKVRDRYKYVLD-DAS 65
XX Db 10 fdwapeagvtl--laggeryemrrpgngpadegwtaadaptdgadvdygylldgei 67
XX QY 66 EIPDPASRYQEGVHGPGSQIIOESKEFNFTFLKEDL---IIEIHVGFTTPEGTFEGV 122
XX Db 68 pldprrtrqpegvhalsrtfdpgahrwqdgwggrelgsgvylhelhigtftpegtldaa 127
XX QY 123 IRKLDYLKDLGTAIEIMPTAQPPGRDWDGYGVYLYAVONSYGSGPGPFKLVDAAHKG 182
XX Db 128 agkldylaglgidfiellpvnafngthwgydgwqfahvhegyggaayqrfvdaahaag 187
XX QY 183 LGVILDVYVNHVPEGNYMVKLGPFYSQKYKTPWGLTFNFDAAESDEVRKFILENVEYWI 242
XX Db 188 lgvldvvyvnhlgpsgnylprygyplkhgegnwtgdsvnldgpgsdhvrqyildnvamw1 247
XX QY 243 KEYNVDGFRDLDAVHAIDTSPKHILEE---IADVWHKYN---RIVIAESDLNDPRVVPNK 296
XX Db 248 rdyrvdgrlridavhaikderavhileefgaladalsesegrpitliaesdnnpriylpr 307
XX QY 297 EKGYNIDAQWDDFHSHIAYLTGERQGYTDFGNLDDIVKSKDVYVDGYKYNFRRK 356
XX Db 308 dvngyglagqwsddfhavhvnvsgettgyysdfslgalakvlrdgffhdgysysfrg 367
XX QY 357 THCEPVELDGNF-----VYIQNDHQVGNRGKGERIILVDRESYKIAAALYLLS 408
XX Db 368 chrpi-----nfsavhpaalvvcsgnhdqignratgdrisqslpygsialaavltg 421
XX QY 409 PYIPMIFMGEGEENPFYFSDSKLIQGVREGKKKE---NGOD-----TDPQDESTF 461
XX Db 422 pftpmifmgeeyatpwwqftshpepelgkatakagrirefermgwdpavvpdpqpetf 481
XX QY 462 NAKSLSKW-----IDEIFSYKILTKMRK-----ELSTACDRRVNVNGENWLIKGR 510
XX Db 482 trskldwaesasagdarlllelysltlrrstpelarlga-dtavefdddarwl----- 535
XX QY 511 EYFS-----LYVFSKSIEVKYSGTLLL 533
XX Db 536 rywrgvqvvinfadrpisldrptall 563
```

```
RESULT 9
AAR77470
ID AAR77470 standard; Protein; 596 AA.
XX
XX AC AAR77470;
XX
XX DT 25-JAN-1996 (first entry)
XX
XX DE Trehalose releasing enzyme.
XX
XX KW Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
XX KW alpha-maltotriosyltrehalose; alpha-maltotetrasyltrehalose;
XX KW maltopentaosyltrehalose; sweetener; taste-improving agent;
XX KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
XX KW pharmaceuticals.
XX
XX OS Rhizobium sp. M11.
XX
XX PN EP671470-A2.
XX
XX PD 13-SEP-1995.
XX
XX PF 07-MAR-1995; 95EP-0301474.
XX
XX PR 07-MAR-1994; 94JP-0059840.
XX
XX PR 07-MAR-1994; 94JP-0059834.
XX
XX PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX PI Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX
XX DR WPI: 1995-312772/41.
XX DR N-PSDB; AAQ98671.
XX
XX PT DNA encoding a trehalose releasing enzyme - which releases trehalose
XX from a non-reducing saccharide having a trehalose structure as an
XX end unit.
XX
XX PS Claim 3; Page 27-29; 45pp; English.
XX
XX CC This enzyme can be used for the preparation of trehalose with high
XX yields and efficiency from non-reducing saccharides such as
XX alpha-glucosyltrehalose, alpha-maltosyltrehalose,
XX alpha-maltotriosyltrehalose, alpha-maltotetrasyltrehalose and
XX maltopentaosyltrehalose. The trehalose can be used as a sweetener,
XX taste-improving agent, quality-improving agent, stabiliser, filler,
XX excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
XX SQ Sequence 596 AA;
XX
XX
XX Query Match 34.4%; Score 1025; DB 16; Length 596;
XX Best Local Similarity 40.3%; Pred. NO. 1.3e-75;
XX Matches 229; Conservative 80; Mismatches 197; Indels 62; Gaps 16;
XX
XX QY 14 FTLWAPYQKSVKLVLEKG-LYEMER-----DEKGYFTITLNNVKVRDRYKYVLD-DAS 65
XX Db 17 fdwapeagvtl--laggeryemrrpgngpadegwtaadaptdgadvdygylldgei 74
XX QY 66 EIPDPASRYQEGVHGPGSQIIOESKEFNFTFLKEDL---IIEIHVGFTTPEGTFEGV 122
XX Db 75 pldprrtrqpegvhalsrtfdpgahrwqdgwggrelgsgvylhelhigtftpegtldaa 134
XX QY 123 IRKLDYLKDLGTAIEIMPTAQPPGRDWDGYGVYLYAVONSYGSGPGPFKLVDAAHKG 182
XX Db 135 agkldylaglgidfiellpvnafngthwgydgwqfahvhegyggaayqrfvdaahaag 194
XX QY 183 LGVILDVYVNHVPEGNYMVKLGPFYSQKYKTPWGLTFNFDAAESDEVRKFILENVEYWI 242
XX Db 195 lgvldvvyvnhlgpsgnylprygyplkhgegnwtgdsvnldgpgsdhvrqyildnvamw1 254
XX QY 243 KEYNVDGFRDLDAVHAIDTSPKHILEE---IADVWHKYN---RIVIAESDLNDPRVVPNK 296
```

Db 255 rdyrvdgirldavhaikderavhileefgaladalssesgrrpltiliaesdlnmprillypr 314
Qy 297 EKCQGVINDAQWDDPHHSIHAAYLGERQGYTDFGNLDDIVKSYKDVYDQKYSNFRK 356
Db 315 dvngyglagwsddfhavhvnvsgettgyysdfdsigalakvirdgffhdgssysfgr 374
Qy 357 THGEPVGLDGCNF-----VVIQNHQDVGNRGKGERIKLVDRSYKIAAALYLLS 408
Db 375 chgrpi-----nfsavhpaalvcvsnhdqignratgdrisqslpygsalaaavltitg 428
Qy 409 PYIMFNGEYGEENPYFFSDFSLSKLIQVREGKKE---NGQD-----TDPQDESTF 461
Db 429 pftmilmgeeygattpwqfftshepeelgkataegrirefermgwdpavvvpdpqdbetf 488
Qy 462 NASLSWK-----IDEEIFSYKILIKMRK-----ELSIACDRRVNVNGENWLIIGKR 510
Db 489 trskldwaasagdhallerylsitlrstpelarlgfa-dtavefdddarwl----- 542
Qy 511 EYFS-----LVFSSKSIEVKYSGTLLL 533
Db 543 rywggqvvinfadrpisldrptall 570
RESULT 10
ID AAY85157 standard; protein; 575 AA.
XX AC AAY85157;
XX DT 23-JUN-2000 (first entry)
XX DE Trehalose-releasing enzyme amino acid sequence.
XX KW Non-reducing saccharide-forming enzyme; trehalose; disaccharide; food;
XX KW high yield; trehalose releasing enzyme; moisture retention; cosmetic;
XX KW production.
XX OS Arthrobacter sp. S34.
XX PN EP990704-A2.
XX PD 05-APR-2000.
XX PF 13-SEP-1999; 99EP-0307220.
XX PR 11-SEP-1998; 98JP-0258394.
XX PR 11-DEC-1998; 98JP-0352252.
XX PR 26-JAN-1999; 99JP-0016931.
XX PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX PI Yamamoto T, Maruta K, Kubota M, Fukuda S, Miyake T;
XX DR WPI; 2000-273225/24.
XX DR N-PSDB; AAA10503.
XX PT New mesophilic enzyme from Arthrobacter, used to convert starch
XX PT hydrolyse into non-reducing sugars, particularly trehalose, useful as
XX PT a sweetener in foods, pharmaceuticals and cosmetics -
XX PS Claim 29; Page 47-48; 93pp; English.

XX This sequence represents a trehalose-releasing enzyme amino acid sequence
XX from Arthrobacter sp. S34. This enzyme specifically hydrolyses a
XX non-reducing saccharide having a trehalose structure as an end unit. The
XX invention relates to this enzyme and also to a non-reducing saccharide
XX forming enzyme that forms a non-reducing saccharide having a trehalose
XX structure as an end unit by reducing partial starch hydrolysates. The two
XX enzymes can be used to produce trehalose, a disaccharide consisting of
XX two moles of glucose bound at their reducing residues. The disaccharide
XX is substantially free of reducibility and has moisture retaining
XX abilities. The enzymes are used to produce non-reducing sugars, they are

CC particularly used for further enzymatic conversion of starch hydrolysate
CC to trehalose. The non-reducing sugars produced by the enzymes are useful
CC in foods, pharmaceuticals and cosmetics. Non-reducing sugars are
CC compatible with amino acids and proteins, they do not cause browning, and
CC have good moisture-retaining properties. The enzymes produce the
CC non-reducing sugars in high yield and function at a medium temperature
CC range i.e. between 40 degrees celsius and 60 degrees celsius, and in an
CC acidic pH range, i.e. a pH of less than 7. Production of non-reducing
CC sugars using the enzymes eliminates the need to convert a starch
CC hydrolysate into sugar alcohols (non-reducing) by catalytic
XX hydrogenation.
XX SQ Sequence 575 AA;

Query Match 34.4%; Score 1024.5; DB 21; Length 575;
Best Local Similarity 39.4%; Pred. No. 1.4e-75;
Matches 215; Conservative 90; Mismatches 189; Indels 51; Gaps 13;

Qy 14 FTLWAPYQKSYKLVLEKGLYE--MERDEKGYFTITL---NNVKRDYRYVLDASEIP 68
Db 5 fpwapaqaavtl-vvgggraelpltrdengwalgqpwgdpdlvd-yylvdgkqpf 62
Qy 69 DPASRYQPEGVHGPSQIIQESKEFNFTLKKED-----LIIYEHVGFPTPGTF 119
Db 63 dprslrqprgvh-----elgrefdparyawgddgrldtgaviyellhvgtftpegtl 116
Qy 120 EGVIKLDYLKLGITAITIEMPIAQFPKRDWGYVLYAVQNSYGGPGRKLVDEAH 179
Db 117 dsairridhvirigvdavellpvnafngthgwgydgvlywavyahpygggyayqfvdach 176
Qy 180 KGLGVILDVVYHNVGPEGNYMVKLGYFSQYKTPMGLTFNFDDESDERVKFLENVE 239
Db 177 arglavqvdyvnhlgpsgnhlpdfigpylgsgaantwgdalnldgplsddevrlyldnav 236
Qy 240 YWKEYNVDFRLDVAHAIIDTSPKHILEIA-----DVVHKYNR--IIVIAESDLNDRPV 293
Db 237 ywirdmhadgliridavhalrdaralhlleelaarvdelagelgrpltiliaesdlnpkli 296
Qy 294 NPKEKCYNIDAOVWDDFHHSIHAYLTERQGYTDFGNLDDIVKSYKDVYDQKYSNF 353
Db 297 rsraahgygidaqwdvhhavhanvctevgyyadfgglavkvfgrgfhdtgwtssf 356
Qy 354 RRKTHGEPVG-ELDGCNFFVYIYNHDOVGNRGKGERIKLVDRSYKIAAALYLLSPYIP 412
Db 357 rerhgrpldpdipfrllvafaqhdqgnravgdmsaqvgegslaaaalvllgftp 416
Qy 413 MIPWGEYGEENPYFFSDFSLSKLIQVREGKKE---NGQD-----TDPQDESTFNASK 465
Db 417 mlfmgeewgartpwqfftshepeelgeatargriaefarmgwdpavvvpdpqdpattfarsh 476
Qy 466 LSWKIDEE-----IFSPYKILIKMRKELSIACDRRVNVNGENWLIIGKREYFSLYVFSK 520
Db 477 ldwseperphagllafytdlialrrelpvdpapare-----vdeadeargvfafsr 526
Qy 521 SSIEV 525
Db 527 gplrv 531

RESULT 11
AAY85165
ID AAY85165 standard; Protein; 575 AA.

XX AC AAY85165;

XX DT 23-JUN-2000 (first entry)

XX DE Trehalose-releasing enzyme amino acid sequence.

XX KW Non-reducing saccharide-forming enzyme; trehalose; disaccharide; food;
XX KW high yield; trehalose releasing enzyme; moisture retention; cosmetic;
XX KW production.


```
Db 22 syelfgahvinegkvgrfcvapharevrlvgsfndwdgtdfrelekvndegvwtivvp 81
Qy 50 NVKVRDRYK--VLDASEI--PDPASRYQPEGVHGPQIIQESK--EFNNETFLKK--- 100
Db 82 enleghlykveivtpdgqvlfkadpyafyselrphstas-iaydlkgvqndqgskrrkr 140
Qy 101 -----EDLIIEIHVGTF--TPEGTF-----EGVIRKLDYKDLGTAIEIMPFAQPPGKRD 150
Db 141 kriydgpmviyelhfgskkdggrfytymadelisylvdhgfhthiellpivehpldrs 200
Qy 151 WGYDGVYLYAVQNSYGGPEGRKLVDEAHKGLGVILDVYVNHV--GPEGNYMVKLGP-- 206
Db 201 wgydgtgyavtsrygtphdfmyfvdrcqgagivmdwvpgfhckdahglymfdgapy 260
Qy 207 -YFSQKYKT--PWGLTFNFDASEI--PDPASRYQPEGVHGPQIIQESK--EFNNETFLKK--- 259
Db 261 eyanekdrenyvwg-tanf-dlgkpevrsflisnalfwleyyhidgfrvdamvmywpn 318
Qy 260 -----DTSKPHILEEADVVHVKY--NRIVIAESDLNDPRVYNPKKCGYNIDAQWVDDF 311
Db 319 ndrilyenpyavfvrkineavfaydpnalmiaedstdwkvptaptgyeglgfnykw---- 374
Qy 312 HHSIHYLTGERQGYTDFGNLDDIVKSYKDFVYDGYKSNFRKTHGEPVGEIDGCGNFV 371
Db 375 -----nmgwmndmlk-ymetppye-----rrhvhngvtfsl-----ly 406
Qy 372 VYION-----HDQVGNRGKGERIKLVDRSYKIAAALYLLSPYI-----PMIFMGE 419
Db 407 aysenfilpfshdevvh-gkksllnkmpggyeekf-aqlrilygymahpgkklifmgne 464
Qy 420 YGEENPFYFFSDFSLSKLIQVREGKENGQDTDPQDESTFNASKLSWKIDEEI----- 474
Db 465 -----faqfde-----wkedeldwvlf 482
Qy 475 -FSFYKILIKRKELSIACDRRVNV-----NGENWLLIKGREYFSLYVP 518
Db 483 dfeIhrkmdymkel-iacykrykpfyelhdpdgfgwidvhaeq-sifsf 532

RESULT 13
ID AAR96109 standard; Protein; 652 AA.
AC AAR96109;
DT 20-AUG-1996 (first entry)
DE Starch-branching-enzyme.
KW Starch-branching-enzyme; thermostable; Q-enzyme; Escherichia coli;
KW starch; amylopectin; glucan; cyclodextrin; food; food-additive;
KW adhesive; biodegradable plastic.
OS Bacillus stearothermophilus strain TRBE14 (FERM P-13916).
FH Key Location/Qualifiers
FT Peptide 238..249
FT /note= "Conserved amylase sequence used to construct primer AAT27126"
FT Peptide 309..317
FT /note= "Conserved amylase sequence"
FT Peptide 414..423
FT /note= "Conserved amylase sequence used to construct primer AAT27127"
XX
PN EP710674-A2.
XX
XX 08-MAY-1996.
XX
XX 13-SEP-1995; 95EP-0250222.
XX
XX 31-JUL-1995; 95JP-0195647.
```

```
PR 13-SEP-1994; 94JP-0218554.
XX (EZAK ) EZAKI GLICO CO LTD.
XX Fujii K, Inanaka T, Nakamura H, Okada S, Takaha T;
XX Takata H, Terada Y, Yanase M;
XX WPI; 1996-223407/23.
XX N-PSDB; AAT27128.
XX
XX New glucan with cyclic, branched core and outer branched region -
XX useful in foods, drinks, adhesives, infusion comps. etc., readily
XX soluble in water forming solns. of low viscosity
XX
XX Example A-2; Page 28-31; 50pp; English.
XX
XX This sequence represents a thermostable starch- branching-enzyme
XX (Q-enzyme, EC-2.4.1.18) from Bacillus stearothermophilus TRBE14
XX (FERM P-13916). Conserved regions present in amylases are used to
XX construct primers AAT27126-7, which are used in PCR to amplify a
XX fragment of the gene, which is then used as a probe to isolate a
XX full-length clone in Escherichia coli. The recombinant enzyme may
XX be reacted with starch or amylopectin to produce a glucan with a
XX degree of polymerisation of at least 50, with an inner branched
XX cyclic region and an outer branched region. The product is freely
XX soluble in water, forming a low-viscosity solution, is more
XX resistance to degradation than native starch, does not react with
XX proteins and amino acids, and is degraded to glucose in vivo, so is
XX easily digestible. The glucan may be used in foods, beverages,
XX food-additives, adhesives, in cyclodextrin production and as a
XX starch substitute in biodegradable plastic.
XX
XX Sequence 652 AA;
```

Query Match 13.4%; Score 399; DB 17; Length 652;

Best Local Similarity 25.2%; Pred. No. 4.2e-24;

Matches 149; Conservative 96; Mismatches 189; Indels 158; Gaps 31;

```
Qy 4 AYKIDGNEVI-----FTLWAPYOKSVKL---KVLEKGLYEMER-DESKGYFTITLN 49
Db 22 syelfgahvinegkvgrfcvapharevrlvgsfndwdgtdfrelekvndegvwtivvp 81
Qy 50 NVKVRDRYK--VLDASEI--PDPASRYQPEGVHGPQIIQESK--EFNNETFLKK--- 100
Db 82 enleghlykveivtpdgqvlfkadpyafyselrphstas-iaydlkgvqndqgskrrkr 140
Qy 101 -----EDLIIEIHVGTF--TPEGTF-----EGVIRKLDYKDLGTAIEIMPFAQPPGKRD 150
Db 141 kriydgpmviyelhfgskkdggrfytymadelisylvdhgfhthiellpivehpldrs 200
Qy 151 WGYDGVYLYAVQNSYGGPEGRKLVDEAHKGLGVILDVYVNHV--GPEGNYMVKLGP-- 206
Db 201 wgydgtgyavtsrygtphdfmyfvdrcqgagivmdwvpgfhckdahglymfdgapy 260
Qy 207 -YFSQKYKT--PWGLTFNFDASEI--PDPASRYQPEGVHGPQIIQESK--EFNNETFLKK--- 259
Db 261 eyanekdrenyvwg-tanf-dlgkpevrsflisnalfwleyyhidgfrvdamvmywpn 318
Qy 260 -----DTSKPHILEEADVVHVKY--NRIVIAESDLNDPRVYNPKKCGYNIDAQWVDDF 311
Db 319 ndrilyenpyavfvrkineavfaydpnalmiaedstdwkvptaptgyeglgfnykw---- 374
Qy 312 HHSIHYLTGERQGYTDFGNLDDIVKSYKDFVYDGYKSNFRKTHGEPVGEIDGCGNFV 371
Db 375 -----nmgwmndmlk-ymetppye-----rrhvhngvtfsl-----ly 406
Qy 372 VYION-----HDQVGNRGKGERIKLVDRSYKIAAALYLLSPYI-----PMIFMGE 419
Db 407 aysenfilpfshdevvh-gkksllnkmpggyeekf-aqlrilygymahpgkklifmgne 464
Qy 420 YGEENPFYFFSDFSLSKLIQVREGKENGQDTDPQDESTFNASKLSWKIDEEI----- 474
Db 465 -----faqfde-----wkedeldwvlf 482
```


Db 465 -----faqfde-----wkfedeldwvlf 482

QY 475 -FSFYKILIKMRKELSIACDRRVNV-----NGENWLIKGRYFSLYVF 518

Db 483 dfehrkmndymkel-lacykrykpfyeldhpggfewidvhnacq-sifsf 532

RESULT 14

AAAP94635

ID AAP94635 standard; protein; 648 AA.

AC AAP94635;

XX

XX

XX 29-JUN-1990 (first entry)

DT

DE

XX B. thuringiensis alpha-amylase.

XX

XX Alpha-amylase; silage; probiotic; lactic acid bacteria;

KW Lactobacillus plantarum; E. faecalis;

KW Bacillus thuringiensis strain H1.1 serotype H1 var thuringiensis.

XX

XX Bacillus thuringiensis.

XX

XX WO8901970-A.

XX

XX 09-MAR-1989.

PD

XX

XX 02-SEP-1988; 88WO-EP00813.

PF

XX

XX 02-SEP-1987; 87EP-0401972.

PR

XX

XX (PLAN-) PLANT GENETIC SYST (UYLO-).

PA

XX

XX Michiels F, Delcour J, Mahillon J, Joos H, Platteeuw C, Jossion K;

PI

XX WPI; 1989-085544/11.

XX

DR N-PSDB; AAN91246.

DR

XX

XX Transformed lactic acid bacteria -

PT using exogenous DNA high codes enzyme that breaks down

PT oligosaccharide and/or polysaccharide, used in silage or as a

PT probiotic

XX

XX Fig 2; pages 2/35-3/35; 112pp; English.

PS

XX

XX The inventors claim an inoculum for silage which comprises lactic

CC acid bacteria, esp. Lactobacillus plantarum and E. faecalis, transformed

CC with an exogenous DNA encoding an enzyme which breaks down an oligosac-

CC -haride and/or a polysaccharide into a fermentable carbohydrate. The

CC exogenous enzyme is especially an alpha-amylase (AAP94635) encoded by

CC the DNA sequence in AAN91246 (claims 2, 6 and 10). The transformed lactic

CC acid bacteria can break down polysaccharides and oligosaccharides in a

CC silage crop to provide a complete source of the fermentable carbohydrates

CC that the lactic acid bacteria need for producing lactic acid. Also

CC claimed is a probiotic characterised by the transformed lactic acid

CC bacteria which can be used for establishing and maintaining optimal

CC intestinal flora in animals, including humans.

XX

SQ Sequence 648 AA;

Query Match 12.9%; Score 385.5; DB 10; Length 648;

Best Local Similarity 23.7%; Pred. No. 5.3e-23;

Matches 149; Conservative 117; Mismatches 231; Indels 131; Gaps 29;

QY 4 AYKIDNEVI-----FTLWAPYQKSVK-LKVLEKGLYEMER-----DERGYFTITLN 49

Db 25 synfaghvvtediegvrftwaphakamsvvgdfnewdyeqhkmqlvteegiwalifip 84

QY 50 NKVDRYKVVLD-----DASEIPDPASRQPEGVHGPSQIIQ-ESKEFNNEF----- 97

Db 85 hieegeikyketlagdvilkadpyavayevrpntasvfdikgyewndknwrkdkkk 144

QY 98 -LKKDLIIYEHVGTF--TPEGTF-----EGVIRKLDYLDKLGITAITAIMPAAQFPGKRD 150

Db 145 siykeamtvyelhfsgwkkkedgtlysermveelipyyvehqfthieimplvehpydys 204

QY 151 WGYDGYLYAVQNSYSGPEGFRKLVDEAHKKGLGVILDVYVNHV--GPEGNYAVKLGPF 208

Db 205 wgyqgtgyaatsrfgtdhmbfvdckhkygigvildwvpgfhckdahglylfdgtpy 264

QY 209 SOKYK-----TPWGLTFNFDAESDEVRKFIENVEYWIKEYNVVDGFRDLDAVHAII--- 259

Db 265 qykdqvpqenpvwg-tvnfdlgr-evrnfliisnalfwmryfhiidfrvdavanmlywnk 322

QY 260 ---DTPSKH---ILEETADVVKHNR--IVIAESDLNDPRVWNPCKCGYNIDAQWDDF 311

Db 323 egqeqsnehavsfrelneavfaedeflmtaeadstawplvtvtptyegglgfnkykw--- 378

QY 312 HHSIHAYLTGEROYYTDFGNLDDIVKSKVDVYDGKYNFPRKTHGEPVCEL---DGC 368

Db 379 -----nmgmndvfk-----ymecapeyirkhihekmfsllyayse 414

QY 369 NFVYIQNHDOVGNRGKGERIIKLVDRSEYKIAAALYLLSPYI-----PMIFMGEEYGE 422

Db 415 nfilpi-shdevvh-gk-ksllnkmppgydwkfaqlrlygyffthpgekillmggefgq 471

QY 423 ENPFYFSDFS-SKLIQVREGKKENGQDTPDQDESTFNASKLSWKID----- 471

Db 472 -----fdewkdledidwnlhdemhymhdyfnelialykrskplwqldhspegfqwid 525

QY 472 -----EEIFSYKILIKMRKELSIACDRRVNVVNGENWLI--IKGREYFSLYVFSKSSIEV 525

Db 526 annedqsfirgqgkedalvvcn--ftkatyenykvpgvdpfeyneilnsdae--- 580

QY 526 KYSGT-----LLSSNNSFPQHIE 544

Db 581 qyggsgqvnkkrllktilpeynqahve 608

RESULT 15

AAAR11271

ID AAR11271 standard; Protein; 639 AA.

XX

XX AAR11271;

XX

XX 29-MAY-1991 (first entry)

DT

XX

XX B.steartothermophilus branching enzyme.

DE

XX

XX Branching enzyme; amylaceous substances; transferase;

KW thermostability.

KW

XX

XX Bacillus steartothermophilus.

OS

XX

XX EP418945-A.

PN

XX

XX 27-MAR-1991.

PD

XX

XX 22-AUG-1990; 90EP-0202264.

PF

XX

XX 23-AUG-1989; 89NL-0002128.

PR

XX

XX (COOP-) COOP VERKOOP AVEBE.

PA

XX

XX Kiel JAKW, Boels JM, Beldman G, Venema G;

PI

XX

XX WPI; 1991-088587/13.

DR

XX

XX N-PSDB; AAQ11072.

DR

XX

XX Branching enzyme obtd. from Bacillus steartothermophilus - has good

PT thermostability, for use in modifying starch-like substances

PT

XX

XX Disclosure; fig 4; 13pp; English.

PS

XX

XX This branching enzyme is derived from B.steartothermophilus and is

CC

CC encoded by the glgB gene. It is stable at high temps. and can be
CC produced in large amts. using recombinant techniques. It is a
CC transferase allowing formation of alpha(1-6) branches in starch-like
CC substances. It is useful for extending the shelf life of such
CC materials and can be used in the prepn. of human and animal foods.
XX

SQ Sequence 639 AA;

```
Query Match          12.68; Score 374.5; DB 12; Length 639;
Best Local Similarity 23.68; Pred. No. 4.2e-22;
Matches 141; Conservative 99; Mismatches 190; Indels 167; Gaps 29;

Qy 4 AYKIDGNEVI-----FTLWAPIYQKSVKL-----KVLEKGLYEMERD 39
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 22 syelfgahvikngmvgtrfcvapharevrlvgsfnewngtnfnlkmkvsngqvmifip 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 40 E-KGYFTITLNNVKVRDRYKYL--DDASEI--PDPAARYQPEGVGPSQIIQ-ESKEF 92
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 82 enleghl-----ykeittndgnvllksdpvafyselrphrtasivynikgyqw 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 93 NNEIFLKK-----EDLIIEIYHVGTF--TPEGTF---EGVIRKLDYLKDLGITALE 138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 ndqtwrrkkrlydgpflfiyehfsgwkkkedgsftyqemaeeilpyvlehgfthie 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 139 IMPIAQFPGRDWDGYLYLAVQNSYGGPEGFRKLVDKDEAHKGLGVILDVYVNHV--GP 196
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 llplvehpfdrsyqggyysatsrgtphdlmyfidrchqagigvildwvpgfhckds 249
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 197 EGNWVKLGPF-----SQYKTPWGLTFNDDAESDEVKFEILENVEYWIKEYNVGFR 251
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 250 hglymfdgapayeyanmqdrenyvwg-tanf-dlgkpevrsflisnalfmweyfhvdgfr 307
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 252 LDVAVHAI-----DTSPKHILREIADVHKY--NRIVIAESDLNDPRVVPKKEC 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 308 vdavannmlywnpsdvlykntyavelfqkinetvfyaydpnlimiaedstdwprvtapydg 367
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 300 GYNIDAQWVDDFHSHIAYLTGERQGYTDFGNLDDIVKSKYKDVYVDGKYSNFRRTTHG 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 glgfnykw-----nmngwmndil-----tymetppehrkyvhn 399
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 360 EPGELGCGNFVYIQN-----HDQVGNRCKGERIIKLVDRSYKTAAL-----YLLS- 408
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 400 kvtfsl----lyaysenfilpfshdevvh-gkksllakmpotyeeekfaqlrlygyllth 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 409 PYIPMIFMGEGYGEENPFYFPFSDSKLIQGVREGRKKENGQDTPQDESTFNASKLSW 468
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 455 pgkklfmggefgq-----fdewkd-----leqldw 480
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 469 KIDEEIFSFYKILIKMRKELSIACDRRVNVV-----NGENWLIIRKREYFSLYVF 518
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 481 mlfd--fdmhrnmmyvkel-lkcykcykplyeldhspdgfewidvhnraq-sifsf 533
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: July 15, 2002, 12:17:32
Job time: 69 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:16:28 ; Search time 25.09 Seconds
(without alignments)
544.197 Million cell updates/sec

Title: US-09-298-924-6
Perfect score: 2978
Sequence: 1 MTFAYKIDGNEVFTLWAPY.....PQHIEGKYFDKGFALYKL 559

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2.6/prodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2.6/prodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2.6/prodata/2/1aa/6A_COMB.pep:*
- 4: /cgn2.6/prodata/2/1aa/6B_COMB.pep:*
- 5: /cgn2.6/prodata/2/1aa/FCUS_COMB.pep:*
- 6: /cgn2.6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2978	100.0	559	US-09-242-690A-15	Sequence 15, Appl
2	1748.5	58.7	556	US-08-505-377-1	Sequence 1, Appl
3	1748.5	58.7	556	US-08-798-269-1	Sequence 1, Appl
4	1748.5	58.7	556	US-09-055-210-1	Sequence 1, Appl
5	1060.5	35.6	598	US-08-399-646-14	Sequence 14, Appl
6	1060.5	35.6	598	US-08-607-321-14	Sequence 14, Appl
7	1060.5	35.6	598	US-08-961-240-14	Sequence 14, Appl
8	1060.5	35.6	598	US-08-605-501-14	Sequence 14, Appl
9	1059	35.6	597	US-08-399-646-4	Sequence 4, Appl
10	1059	35.6	597	US-08-607-321-4	Sequence 4, Appl
11	1059	35.6	597	US-08-961-240-4	Sequence 4, Appl
12	1059	35.6	597	US-08-605-501-4	Sequence 4, Appl
13	1025	34.4	589	US-08-399-646-2	Sequence 2, Appl
14	1025	34.4	589	US-08-607-321-2	Sequence 2, Appl
15	1025	34.4	589	US-08-961-240-2	Sequence 2, Appl
16	1025	34.4	589	US-08-605-501-2	Sequence 2, Appl
17	1025	34.4	596	US-08-399-646-12	Sequence 12, Appl
18	1025	34.4	596	US-08-607-321-12	Sequence 12, Appl
19	1025	34.4	596	US-08-961-240-12	Sequence 12, Appl
20	1025	34.4	596	US-08-605-501-12	Sequence 12, Appl
21	399	13.4	652	US-08-528-026C-4	Sequence 4, Appl
22	346.5	11.6	713	US-09-346-237-9	Sequence 9, Appl
23	343	11.5	718	US-09-346-237-10	Sequence 10, Appl
24	334	11.2	606	US-09-187-124-2	Sequence 2, Appl
25	331.5	11.1	893	US-09-514-302-4	Sequence 4, Appl
26	331.5	11.1	1938	US-09-514-302-2	Sequence 2, Appl
27	329.5	11.1	818	US-08-410-784A-2	Sequence 2, Appl

28	329.5	11.1	818	US-09-346-237-11	Sequence 11, Appl
29	321.5	10.8	740	US-08-410-784A-4	Sequence 4, Appl
30	303.5	10.2	829	US-09-514-599-6	Sequence 6, Appl
31	303.5	10.2	862	US-09-346-237-1	Sequence 1, Appl
32	295	9.9	555	US-08-039-777-3	Sequence 3, Appl
33	293.5	9.9	555	US-08-611-361A-3	Sequence 3, Appl
34	293.5	9.9	921	US-09-514-599-2	Sequence 2, Appl
35	291	9.8	915	US-09-346-237-2	Sequence 2, Appl
36	291	9.8	928	US-09-514-599-4	Sequence 4, Appl
37	286	9.6	928	US-08-474-140-11	Sequence 11, Appl
38	286	9.6	928	US-08-477-630-11	Sequence 11, Appl
39	286	9.6	928	US-08-472-293-11	Sequence 11, Appl
40	286	9.6	928	US-08-474-545-11	Sequence 11, Appl
41	286	9.6	928	US-08-478-341-11	Sequence 11, Appl
42	286	9.6	928	US-08-996-733-11	Sequence 11, Appl
43	285	9.6	963	US-08-537-002A-3	Sequence 3, Appl
44	285	9.6	963	US-08-863-010-3	Sequence 3, Appl
45	285	9.6	963	US-09-024-429-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-09-242-690A-15
; Sequence 15, Application US/09242690A
; Patent No. 6284534

GENERAL INFORMATION:

APPLICANT: KONDO, KEIJI

INVENTOR: MIURA, YUTAKA

TITLE OF INVENTION: YEAST VECTOR AND METHOD OF PRODUCING PROTEINS USING THE

FILE OF INVENTION: SAME

FILE REFERENCE: 049441/0118

CURRENT APPLICATION NUMBER: US/09/242,690A

PRIOR FILING DATE: 1999-02-23

PRIOR APPLICATION NUMBER: PCT/JP97/02924

PRIOR FILING DATE: 1997-08-22

PRIOR APPLICATION NUMBER: JP 8/241062

NUMBER OF SEQ ID NOS: 66

SOFTWARE: Patent Ver. 2.1

SEQ ID NO 15

LENGTH: 559

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein
US-09-242-690A-15

Query Match 100.0%; Score 2978; DB 4; Length 559;

Best Local Similarity 100.0%; Pred. No. 4.1e-250; Indels 0; Gaps 0;

Matches, 559; Conservative 0; Mismatches 0;

QY	1	MTFAYKIDGNEVFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYV	60
Db	1	MTFAYKIDGNEVFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYV	60
QY	61	LDASEIPDPASRYQPEGVHGPSQIIQESKEFNFTFLKEDLIYIEHVGTTPGTFEE	120
Db	61	LDASEIPDPASRYQPEGVHGPSQIIQESKEFNFTFLKEDLIYIEHVGTTPGTFEE	120
QY	121	GVTRKLDYKLDGITAIEIMPITQAQFPGKRDWGYDGYLVAVQNSYGGPGEFRKLVDENK	180
Db	121	GVTRKLDYKLDGITAIEIMPITQAQFPGKRDWGYDGYLVAVQNSYGGPGEFRKLVDENK	180
QY	181	KGLGVILDVVYVNHVGPENYMWKLGYPFSQYKTPMGLTFNFDDEADEVKFEILENVEY	240
Db	181	KGLGVILDVVYVNHVGPENYMWKLGYPFSQYKTPMGLTFNFDDEADEVKFEILENVEY	240
QY	241	WIKEYNVDFRLDAVHAIDTSPKHILEETADVHHYKNRIVIAESDLNDPRVYNPKKCG	300
Db	241	WIKEYNVDFRLDAVHAIDTSPKHILEETADVHHYKNRIVIAESDLNDPRVYNPKKCG	300

102(e) date for
this patent is 2/23/99
Senior to mine

QY 301 YNIDAQWDDFHHSIHAYLTGERGYTDFGNLDDIVKSKDVFDYDGKYSNFRKTHGE 360
Db 301 YNIDAQWDDFHHSIHAYLTGERGYTDFGNLDDIVKSKDVFDYDGKYSNFRKTHGE 360
QY 361 PVGELDCGNFVYIQNDHQVGNRGKGRIRIKLVDRYSYKIAAALYLLSPYIPMFMGEEY 420
Db 361 PVGELDCGNFVYIQNDHQVGNRGKGRIRIKLVDRYSYKIAAALYLLSPYIPMFMGEEY 420
QY 421 GEENPFYFSDSKLIQVREGKRGKENGODTDPQDESTFNASKLSWKIDEEIFSFKI 480
Db 421 GEENPFYFSDSKLIQVREGKRGKENGODTDPQDESTFNASKLSWKIDEEIFSFKI 480
QY 481 LIKMRKELSIACDRRVNVNGENWLIIGREYFSLYVFSKSSIEVKYSYGTILLSSNNSFP 540
Db 481 LIKMRKELSIACDRRVNVNGENWLIIGREYFSLYVFSKSSIEVKYSYGTILLSSNNSFP 540
QY 541 QHIEGKYFEDKGFALYKL 559
Db 541 QHIEGKYFEDKGFALYKL 559
RESULT 2
US-08-505-377-1
; Sequence 1, Application US/08505377
; Patent No. 5856146
; GENERAL INFORMATION:
; APPLICANT: MITSUZUMI, Hitoshi
; APPLICANT: KUBOTA, Michio
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME WHICH
; OPERATING SYSTEM: PC-DOS/MS-DOS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,377
; FILING DATE: 21-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 190180/1994
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 109128/1995
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP NOT YET RECEIVED
; FILING DATE: 04-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: MITSUZUMI-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; TELEX: 249688
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-505-377-1

Query Match 58.7%; Score 1748.5; DB 2; Length 556;
Best Local Similarity 58.9%; Pred. No. 1.9e-143;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;
QY 2 TFAYKIDGNEVIFTLWAPYQSKVLKYLEKGLYEMERDEKGYFTITLNNKVRDRYKYL 61
Db 3 SFGNIEKNGKIFLWAPYVNSVKL-LSKKLIPMERNDGFEVEIDDEENLTYSYII 61
QY 62 DDASEIPDPSRQPEGVHGPSQIIQESKEFNNETFLKKEDLIYIYHVGTFPEGTPEG 121
Db 62 EDKEIPDPSRQPLGVHDKSQLRTDYQILDGLGVKVEDLIYIYHVGTFPEGTPEG 121
QY 122 VIRKLDYKLDGTAIEIMPIAOPPGKRDGCVLYAVQNSYGGEGEGKLVDEAHKK 181
Db 122 VIRKLDYKLDGTAIEIMPIAOPPGKRDGCVLYAVQNSYGGEGEGKLVDEAHKK 181
QY 182 GLGVLDVYVNHVPEGNVMKLGYPFSQYKTPWGLTFNFDADAESDEVKRFLENYEYW 241
Db 182 GIAVILDVYVNHVPEGNVLLGLGYPFSQYKTPWGLTFNFDADAESDEVKRFLENYEYW 241
QY 242 IKEYNVGDFRLDAVHAIDTSPKHILEEIAADVHKNYRIVIAESDLNDPRVNPKEKCY 301
Db 242 FKTFKIDGLRLDAVHAIFDNSPKHILQEIIEAKHAGLQGVIAESDLNDPKIV--KDCGY 299
QY 302 NIDAQWDDFHHSIHAYLTGERGYTDFGNLDDIVKSKDVFDYDGKYSNFRKTHGE 361
Db 300 KIDAQWDDFHHSIHAYLTGERGYTDFGNLDDIVKSKDVFDYDGKYSNFRKTHGE 359
QY 362 VGLDGCNFFVYIQNDHQVGNRGKGRIRIKLVDRYSYKIAAALYLLSPYIPMFMGEEY 421
Db 360 VGLDPPKRFVYIQNDHQVGNRGKGRIRIKLVDRYSYKIAAALYLLSPYIPMFMGEEY 419
QY 422 ENPFFYFSDSKLIQVREGKRGKENGODTDPQDESTFNASKLSWKIDEEIFSFKIL 481
Db 420 ETNPFYFSDSKLIQVREGKRGKENGODTDPQDESTFNASKLSWKIDEEIFSFKIL 479
QY 482 IKMRKELSIACDRRVNVNGENWLIIGREYFSLYVFSKSSIEVKYSYGTILLSSNNSFP 541
Db 480 INKRYN-NCKRVKEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGI--GFPK 536
QY 542 HIEEGK-YEFDKGFALYKL 559
Db 537 KLKKDELIVKVRGVVYQL 555
RESULT 3
US-08-798-269-1
; Sequence 1, Application US/08798269
; Patent No. 6027918
; GENERAL INFORMATION:
; APPLICANT: MITSUZUMI, Hitoshi
; APPLICANT: KUBOTA, Michio
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME WHICH
; OPERATING SYSTEM: PC-DOS/MS-DOS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/798,269
; FILING DATE:

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,377
; FILING DATE: 21-JUL-1995
; APPLICATION NUMBER: JP 190180/1994
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 109128/1995
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP NOT YET RECEIVED
; FILING DATE: 04-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: MITSUZUMI-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 556 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-798-269-1

Query Match      58.7%; Score 1748.5; DB 3; Length 556;
Best Local Similarity 58.9%; Pred. No. 1.9e-143;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

QY 2 TFAYKIDGNEVFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITNNVKVRDRYKVL 61
DB 3 SFGGNEKKGIFKLWAPYVNSVKL-LSKKLIPMEKDEGFEVEIDDEENTYSYII 61
QY 62 DDASEIPDPASRYOPEGVHGSPQIIOESKEFNFTFLKEDLIIEHVGTFTPEGFEG 121
DB 62 EDKREIPDPASRYOPLGVHDKSQLIRTDYQILDGLKVKIEDLIIEHVGTFTPEGFEG 121
QY 122 VIRKLDYLKDLGITAIEMPTAQPGKRDGVDGYVLYAVONSGYGGEGFKLVDEAHKK 181
DB 122 VIEKLDYLKDLGITGIELMPVAQPGNDRDGYDGVLYAVQNTYGGVWELAKLVNEAHR 181
QY 182 GLGVILDVYVNHGPEGNYMVKLGYPFSQKYKTPWGLTFNFDAAESDEVKRFILNEYW 241
DB 182 GIAVILDVYVNHGPEGNYMVKLGYPFSQKYKTPWGLTFNFDAAESDEVKRFILNEYW 241

Query Match      58.7%; Score 1748.5; DB 4; Length 556;
Best Local Similarity 58.9%; Pred. No. 1.9e-143;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

QY 2 TFAYKIDGNEVFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITNNVKVRDRYKVL 61
DB 3 SFGGNEKKGIFKLWAPYVNSVKL-LSKKLIPMEKDEGFEVEIDDEENTYSYII 61
QY 62 DDASEIPDPASRYOPEGVHGSPQIIOESKEFNFTFLKEDLIIEHVGTFTPEGFEG 121
DB 62 EDKREIPDPASRYOPLGVHDKSQLIRTDYQILDGLKVKIEDLIIEHVGTFTPEGFEG 121
QY 122 VIRKLDYLKDLGITAIEMPTAQPGKRDGVDGYVLYAVONSGYGGEGFKLVDEAHKK 181
DB 122 VIEKLDYLKDLGITGIELMPVAQPGNDRDGYDGVLYAVQNTYGGVWELAKLVNEAHR 181
QY 182 GLGVILDVYVNHGPEGNYMVKLGYPFSQKYKTPWGLTFNFDAAESDEVKRFILNEYW 241
DB 182 GIAVILDVYVNHGPEGNYMVKLGYPFSQKYKTPWGLTFNFDAAESDEVKRFILNEYW 241
```



```
QY 221 NFDAAESDEVKRFILNVEYWIKEYNVVDGFRDLDAVHAIIDTSPKHILEIEIADVHVHKYRI 280
Db 235 NLDGPGSDHVRVYILDNLAMWLRDYVDGLRLDAVHALKDERAVHILEDFGALADQISAE 294
QY 281 V-----IAESLDNPRVVPKCCYNYDAQWDDFHSHIAYLTGERQGYTDFGNLD 334
Db 295 VGRPLTLIAESLDNPRLLYPRDVNGYLGEGOWSDDFHHAHVNVVTGETTGYISDFDLSA 354
QY 335 DIVKSYKDVYVYDGKYSNFRKTHGEPVGLDGCNF-----VVYIQNHQDVGNRKG 386
Db 355 ALAKVLRDGFFHDGYSVSSFRERHGRPI-----NFSAVHPAALVVCSONHDIQGNRATG 408
QY 387 ERIIKLVDRSYKIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDSKLIQGVREGRK 446
Db 409 DRLSQTLPYGSIALAAVLTLCGPTPMLLMGEEYGASTPWQFFTSHPPELGKATAEGRI 468
QY 447 KE---NGQD---TDPDESTFNASKLSWKIDEE-----IFSFKILIKMKRSELSIAC-- 492
Db 469 KEFERMGWDPAVVPDQDPETFRRSKLDWAEAAEGDHARLLELYSLTALRRSTPDLTKL 528
QY 493 ---DRRVNVNGENWMLIIKREYFSLYVFSKSSIEVKYSGTLLL 533
Db 529 GFEDTQVAFDEADARWLFRFRGGVQVLLNFSQPVSLDGAGTALL 572

RESULT 8
US-08-605-501-14
; Sequence 14, Application US/08605501
; Patent No. 5834287
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HATTORI, Kazuko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,501
; FILING DATE: 26-FEB-1996
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,646
; FILING DATE: 07-MAR-1995
; APPLICATION NUMBER: JP 59834
; FILING DATE: 07-MAR-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 59840
; FILING DATE: 07-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA=5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 598 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-501-14
```

```
Query Match 35.6%; Score 1060.5; DB 2; Length 598;
Best Local Similarity 39.7%; Pred. No. 1e-83;
Matches 232; Conservative 92; Mismatches 197; Indels 63; Gaps 16;
```

```
QY 1 MTFAYKIDGNEVI-----FTLHAPYQKSVKLVLEKGLYEMER-----DEKGYFTI-- 46
Db 1 MHTHTPREAAKPVLPARYDVMAPNAESVTL--LAGGERYAMORRAETGPDAGWMTAAG 58
QY 47 --TLNNVKVRDRYKVVLD--DASEIPDPASRYQPEGVHGPSIIQESK--EFNETFLKKE- 101
Db 59 APTDGNVD---YGLLDGDETPLPDPTRRQPDGVHSLSRTFDFSAWSQDDANQGRRL 114
QY 102 -DLIYIEIHVGFTEGTFEGVIRKLDYLKLGITAIEIMPIAQPGKRDWGYDGYLYA 160
Db 115 QGAVIYELHLGTFTEGTLLEAAAGKLDYLAGLVDFIELLPVNAFNGTHNMGYDGVQWFA 174
QY 161 VONSYGGEPEGRKLYDEAHKGLGVLDVVYVNHVGPENYVVKLGPYFSQKYKTPWGLTF 220
Db 175 VHEAYGGPEAYQRFVDAHAAGLVQDVVYVNHLPNGNYLPRFGPYLKQEGEWTGDSV 234
QY 221 NFDAAESDEVKRFILNVEYWIKEYNVVDGFRDLDAVHAIIDTSPKHILEIEIADVHVHKYRI 280
Db 235 NLDGPGSDHVRVYILDNLAMWLRDYVDGLRLDAVHALKDERAVHILEDFGALADQISAE 294
QY 281 V-----IAESLDNPRVVPKCCYNYDAQWDDFHSHIAYLTGERQGYTDFGNLD 334
Db 295 VGRPLTLIAESLDNPRLLYPRDVNGYLGEGOWSDDFHHAHVNVVTGETTGYISDFDLSA 354
QY 335 DIVKSYKDVYVYDGKYSNFRKTHGEPVGLDGCNF-----VVYIQNHQDVGNRKG 386
Db 355 ALAKVLRDGFFHDGYSVSSFRERHGRPI-----NFSAVHPAALVVCSONHDIQGNRATG 408
QY 387 ERIIKLVDRSYKIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDSKLIQGVREGRK 446
Db 409 DRLSQTLPYGSIALAAVLTLCGPTPMLLMGEEYGASTPWQFFTSHPPELGKATAEGRI 468
QY 447 KE---NGQD---TDPDESTFNASKLSWKIDEE-----IFSFKILIKMKRSELSIAC-- 492
Db 469 KEFERMGWDPAVVPDQDPETFRRSKLDWAEAAEGDHARLLELYSLTALRRSTPDLTKL 528
QY 493 ---DRRVNVNGENWMLIIKREYFSLYVFSKSSIEVKYSGTLLL 533
Db 529 GFEDTQVAFDEADARWLFRFRGGVQVLLNFSQPVSLDGAGTALL 572
```

RESULT 9

```
US-08-399-646-4
; Sequence 4, Application US/08399646
; Patent No. 5556781
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HATTORI, Kazuko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```


OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/399,646
FILING DATE: 07-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 597 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-399-646-4

Query Match 35.6%; Score 1059; DB 1; Length 597;
Best Local Similarity 40.5%; Pred. No. 1.4e-83;
Matches 229; Conservative 89; Mismatches 190; Indels 58; Gaps 15;

QY 14 FTLWAPYQKSVKLVLEKG-LYEMER-----DEKGYFTI-----TLNNVKVRDRYKYVLD 62
DB 18 IDVWAPNAESVTL--LAGGERYAMORRAETGPDAGWMTAAGAPTDCGNVD-----YGYLLD 71
QY 63 -DASEIPDPASRYQPGVHGPSQIIQESK-EFNNETFLKKE--DLIIYEHVCTFTPEGT 118
DB 72 GDETPLPDPRTRQPGVGHALSRTDFPSAYSWQDDAQWQRELGAGVYIELHLGTFTEPGT 131
QY 119 FEGVIRKLDYKLDGITAIEIMPIAOPGKRDMGYDGVLYAVONSYGGPEGRKLVDEA 178
DB 132 LEAAAGKLDYLAGLVDFIELLEPVNAFNGTHNWGYDGVQWFAVHEAYGGPEAYQRFVDA 191
QY 179 HKKGLGVLDVYVNHVGPENYVMVKGPFYSQKYKTPWGLTFNFDAAESDEVRKFTILENV 238
DB 192 HAAGLVQIDVYVNHVGPENYVMVKGPFYSQKYKTPWGLTFNFDAAESDEVRKFTILENV 238
QY 239 EYWIKEYNDGFRDLAVHAIIDTSPKHILEEADVYVHKNRYV-----IAESDLNDPRV 292
DB 252 AMWLRDVRVGLRLDAVHALKDERAVHILEDFGALADQISAEVGRPLTLIAESDLNPR 311
QY 293 VNPKEKCGYNIDAQWDDFHSTHAYLTGRQGYTDFGNLDDIVKSYKDVYFDGKYSN 352
DB 312 LYPRDNGYGLEQWDDFHSHAVHVNVTGETTGYSDFDLSAALAKLVKRGDFHDSYSS 371
QY 353 FRKTHGPEVGLDGNF-----VYIQNHQDVGNCKGERIKLVDRSYKTAAL 404
DB 372 FRERHGRPL-----NFSVHPAALVVCSONHDQIGNRATGDRLSQTLPGYSLAALV 425
QY 405 YLLSPYIPIMFMEEGEENPFYFFSDFSDSKLIQVREGRKKE---NGQD-----TDPOD 457
DB 426 TLGPTPTMLLMGEEYGASTPWQFFSHPELPGKATAEGRKEFERFMGWDPAVVPDQD 485
QY 458 ESTFNASKLSWKIDEE-----IPSFYKILMKRKELSIAC-----DRRVNVNNGENWLI 507
DB 486 PETFRRSKLDWAAEAGCDHARLELYRSLTALRRSTPDLTKLGFEDTQVAFEDARWLR 545
QY 508 KGREYFSLYVFSKSSIEVYKSTLL 533
DB 546 RRGGVQVLLNFSEQPVSLDAGAGTALL 571

RESULT 10
US-08-607-321-4
Sequence 4, Application US/08607321
Patent No. 5716813
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Keiji
APPLICANT: HATTORI, Kazuko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/607,321
FILING DATE: 26-FEB-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,646
FILING DATE: 07-MAR-1995
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 597 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-607-321-4

Query Match 35.6%; Score 1059; DB 1; Length 597;
Best Local Similarity 40.5%; Pred. No. 1.4e-83;
Matches 229; Conservative 89; Mismatches 190; Indels 58; Gaps 15;

QY 14 FTLWAPYQKSVKLVLEKG-LYEMER-----DEKGYFTI-----TLNNVKVRDRYKYVLD 62
DB 18 IDVWAPNAESVTL--LAGGERYAMORRAETGPDAGWMTAAGAPTDCGNVD-----YGYLLD 71
QY 63 -DASEIPDPASRYQPGVHGPSQIIQESK-EFNNETFLKKE--DLIIYEHVCTFTPEGT 118
DB 72 GDETPLPDPRTRQPGVGHALSRTDFPSAYSWQDDAQWQRELGAGVYIELHLGTFTEPGT 131
QY 119 FEGVIRKLDYKLDGITAIEIMPIAOPGKRDMGYDGVLYAVONSYGGPEGRKLVDEA 178
DB 132 LEAAAGKLDYLAGLVDFIELLEPVNAFNGTHNWGYDGVQWFAVHEAYGGPEAYQRFVDA 191
QY 179 HKKGLGVLDVYVNHVGPENYVMVKGPFYSQKYKTPWGLTFNFDAAESDEVRKFTILENV 238
DB 192 HAAGLVQIDVYVNHVGPENYVMVKGPFYSQKYKTPWGLTFNFDAAESDEVRKFTILENV 238

Db 536 RYWRGGVVVLFNFAADRPISLDRPGTALL 563

Search completed: July 15, 2002, 12:18:04
Job time: 96 sec

```
Query Match      34.4%; Score 1025; DB 2; Length 589;
Best Local Similarity 40.3%; Pred. No. 1.2e-80;
Matches 229; Conservative 80; Mismatches 197; Indels 62; Gaps 16;
```

QY	14	FTLWAPYQKSVKLVLSEKG-LYEKER-----DEKGYFTTLNNVKVRDRKYVLD-DAS	65
DB	10	FDIWAPEAGTVTL--LAGGERYEMRRPGNGPADEGWTAADTAFGADVVDGYLLDGDGEI	67
QY	66	EIPDPASRYOPEGVGHGSQIIQESKEFNNETFLKKEDL---IIVEIHVGFTPPGCTEGV	122
DB	68	PLPDPTTRPOEGVHALSRTPDGAHRWDQAGWGROELQGSVIYELHIGTFTPEGTLDAA	127
QY	123	IRKLDYLKDGLGITAIEIMPTAQQPKCKDWGVLYLVAVQNSYGSGPEGFKVLDEAHKKG	182
DB	128	AGKLDYLAGLGIDFIELLPYNFNGTHNWGYDGQWFVAFVHEGYGPAAQYRFVDAHAAG	187
QY	183	LGVILDVVVNVHVPGEINVMKVLGFYSOKYKTPMGLTFNFDADAESDEVKRPFILENVEWI	242
DB	188	LGVIQDVVVVNHLPSPGNLPRYPGLYKHGECNTMGDSVNLDGPGSDHVQRQILDNVAMWL	247
QY	243	KEYNVDFGRLDVAHAIIDTSPKHTILER---IADVVVHKVN--RTVIAESLNDPRVNPVK	296
DB	248	RDRVGLRLDLNDAHAKDERAVHILLEFFGALADASEGGRPRTLIAESLNNPRLYPR	307
QY	297	EKCIGNIDAQVDDFHHSIIHAYLTGEROGYYTDGFLNDDIVKSKYDVYDGKYSNFRRK	356
DB	308	DVNGYGLAGOWSDDFHHAVHVNVSGETGYYSDFSLSGALAKVLURDGFHFHDGYSYSPGR	367
QY	357	THECPYVELDCGNF-----VVYIONHDQVNRGKGIITKLVDRESYKIAAAYLLS	408
DB	368	CHGRPI-----NFSAVHPAALVCSOHDOIGNRATGRDLSQSLSPTYGSLAALAVLTG	421
QY	409	PYIPMFPMGEYGEENPFYFSDFSCLKIQGVREGKRKE---NCQD----TDPQDESTF	461
DB	422	PFTPLFMGEYGAATTWPQFTTSHPPELGPALKATAEGRIFERFERMGWDPVAVVPDPQDPETF	481
QY	462	NASKLSWK-----IDEIFSFYKILLIKWRK-----ELSTACDRRVNVNGENWLIIGR	510
DB	482	TRSKLDWAESAAGDHARLELILRSLTRSTPELARLGFA-DTAVFEDDDARWL-----	535
QY	511	EYFS-----LVYFSKSISIEVYSGTLLL	533

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:16:28 ; Search time 32.43 Seconds
(without alignments)
1656.304 Million cell updates/sec

Title: US-09-298-924-6
Perfect score: 2978
Sequence: 1 MTFAYKIDGNEVIFTLWAPY.....PQHIEGKYEFDFKGFALYKL 559

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2973	99.8	558	2	JC5135
2	2467	82.8	561	2	S73087
3	1748.5	58.7	556	2	JC5132
4	1288.5	43.3	620	2	AH1827
5	1060.5	35.6	598	2	S65770
6	1025	34.4	596	2	JC4697
7	1024.5	34.4	575	2	JC7727
8	997.5	33.5	580	2	G70763
9	978.5	32.9	594	2	A10673
10	905.5	30.4	600	2	H75516
11	881.5	29.6	583	2	G83375
12	876.5	29.4	601	2	D96001
13	860.5	28.9	595	2	AG3194
14	524.5	17.6	552	2	S19134
15	524.5	17.6	552	2	AH1915
16	391	13.1	666	2	B56639
17	374.5	12.6	627	1	S40048
18	374.5	12.6	638	2	S18599
19	365	12.3	1280	2	E95031
20	364	12.2	1256	2	G97902
21	361	12.1	630	2	D70363
22	349.5	11.7	783	2	B84823
23	348.5	11.7	666	2	E71565
24	347.5	11.7	764	2	AG1895
25	345.5	11.6	422	2	S31839
26	345.5	11.6	664	2	B86539
27	345.5	11.6	664	2	B72084
28	345.5	11.6	666	2	G81717
29	345.5	11.6	707	2	S77094

30	344.5	11.6	741	1	S34218	1,4-alpha-glucan b
31	343	11.5	718	2	I73088	glycogen operon pr
32	340	11.4	730	1	I64118	1,4-alpha-glucan b
33	339.5	11.4	789	2	T01321	probable isoamylas
34	336	11.3	702	2	B85091	isoamylase-like pr
35	335	11.2	764	2	S47569	1,4-alpha-glucan b
36	335	11.2	1072	2	A84112	alkaline amylopull
37	334.5	11.2	717	2	H84057	pullulanase amyl l
38	333.5	11.2	720	2	B75540	glycogen operon pr
39	333.5	11.2	770	2	S76095	hypothetical proce
40	330.5	11.1	562	2	C82693	trehalose-6-phosph
41	329.5	11.1	818	2	T02231	probable isoamylas
42	329	11.0	733	2	T49679	probable branching
43	327.5	11.0	735	2	F98228	1,4-alpha-glucan b
44	327.5	11.0	735	2	AH3057	glycogen branching
45	326.5	11.0	603	2	AC0674	probable glycogen

ALIGNMENTS

RESULT 1

JC5135

alpha-amylase (EC 3.2.1.1) - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000

C:Accession: JC5135

R:Kobayashi, K.; Kato, M.; Miura, Y.; Kettoku, M.; Kameda, T.; Iwamatsu, A.

Biosci. Biotechnol. Biochem. 60, 1882-1885, 1996

A:Title: Gene cloning and expression of new trehalose-producing enzymes from the hype

A:Reference number: JC5134; MUID:97141610

A:Accession: JC5135

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-558 <KOB>

A:Cross-references: DBJ: D64130; NID: g987702; PID: BAA11010.1; PID: g1838936

A:Experimental source: strain KM1

C:Comment: This enzyme hydrolyzes starch alpha-1,4 linkage adjacent to the alpha-1,1

C:Superfamily: trehalose trehalohydrolase

C:Keywords: glycosidase; hydrolase

Query Match 99.8%; Score 2973; DB 2; Length 558;

Best Local Similarity 100.0%; Pred. No. 1.1e-185;

Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TFAYKIDGNEVIFTLWAPYQSKVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYVL 61

DB 1 TFAYKIDGNEVIFTLWAPYQSKVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYVL 60

QY 62 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNEFLKKEDLIIYEHVGTFTPEGTRFG 121

DB 61 DDASEIPDPASRYQPEGVHGPSQIIQESKEFNNEFLKKEDLIIYEHVGTFTPEGTRFG 120

QY 122 VIRKLDYLKDLGTAITEIMPIAQQFPGKRDWGDVYLYAVQNSYSGGEGEGRKLIVDEAHKK 181

DB 121 VIRKLDYLKDLGTAITEIMPIAQQFPGKRDWGDVYLYAVQNSYSGGEGEGRKLIVDEAHKK 180

QY 182 GLGVILDVYVNHVGPBGNTVMVKGPFYSQKYTPGWLTFNFDADAESDEVKRFILNVEYVW 241

DB 181 GLGVILDVYVNHVGPBGNTVMVKGPFYSQKYTPGWLTFNFDADAESDEVKRFILNVEYVW 240

QY 242 IKEYNVDDGFRDLDAVHAIDTSPKHILEETADVVKHKNRIVIAESDLNDRVVPNPKKCY 301

DB 241 IKEYNVDDGFRDLDAVHAIDTSPKHILEETADVVKHKNRIVIAESDLNDRVVPNPKKCY 300

QY 302 NIDAQWVDDFHSHIAYLTGERQGYTDFGNLDDIVKSYKDVYVDGKYSNFRKTHGEP 361

DB 301 NIDAQWVDDFHSHIAYLTGERQGYTDFGNLDDIVKSYKDVYVDGKYSNFRKTHGEP 360

QY 362 VGEIDGCFVYVYIQNHQDQVGNRGKGRRIIKLVDRRESYKIAAALYLLSPYIPMIFMGEYVG 421

DB 361 VGEIDGCFVYVYIQNHQDQVGNRGKGRRIIKLVDRRESYKIAAALYLLSPYIPMIFMGEYVG 420

C:Species: Arthrobacter sp.
A:Variety: strain Q36
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S65770
R:Maruta, K.; Hattori, K.; Nakada, T.; Kubota, M.; Sugimoto, T.; Kurimoto, M.
Biochim. Biophys. Acta 1289, 10-13, 1996
A:Title: Cloning and sequencing of trehalose biosynthesis genes from Arthrobacter sp.
A:Reference number: S65769; MUID:96195835
A:Accession: S65770
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-598 <MAR>
A:Cross-references: EMBL:D63343; NID:g1255444; PIDN:BAA09668.1; PID:g1255446
C:Superfamily: trehalose trehalohydrolase

Query Match 35.6%; Score 1060.5; DB 2; Length 598;
Best Local Similarity 39.7%; Pred. No. 2.7e-61;
Matches 232; Conservative 92; Mismatches 197; Indels 63; Gaps 16;

QY 1 MTFAYKIDGNEVI-----FTLWAPYOKSVKLVLEKG-LYEMER-----DEKGYFTI-- 46
DB 1 MHTYTPREAAKPVLPARVDWAPNAESVTL--LAGGERYAMORRAETGPDAGWWTAAG 58

QY 47 --TLNNVKVRDXYKVVLD-DASEIPDPASRYQPEGVHGSQIIQESK-EFNNETFLKKE- 101
DB 59 APTDGNVD---YGLDGDDETPDPRTTRQPDGVHALSRTFDPSAYSQWDDAQWQREL 114

QY 102 -DLIIYEHVGTFTPEGTGEGVIRKLDYLDLIGITAEIMPIAQPPGKRDGWDGYLYA 160
DB 115 QGAVIYELHGTFTPEGTLEAAAGKLDYLAGLVDFIELLPVNAFNGTHNGYDGYQWPA 174

QY 161 VONSYGPPGFKLVLDDEAHKGLGVLDVYVNHVPEGMVMVKLGPFYSQYKTPWGLTF 220
DB 175 VHEAAGGPAYORFVDAHAAGLVQDVVYVNHLPFGSGNYLPRFGPYLKQEGNTWGDV 234

QY 221 NFDDAESDEVRFKLEINVEYIKEYNVGDFRLDAVHAIIIDTSPKHILEEIAADVVKYNR 280
DB 235 NLDGPGSDHVRVYILDNLAMWLDYRDVGLRLDAVHALKDERAVHILEDFGALADQISAE 294

QY 281 V-----IAESDINDPRVVPKCEGYNIDAOQWDDFHHSIHAYLTGRQGYTDFGNLD 334
DB 295 VGRPLTLIAESDINNPRLLYPRDVGNGYLEGQWSDDFHVAHVNVGTGTYGYSDFDLSA 354

QY 335 DTVKSVKDVVFYDGVKSNFRKTHGPELDCGNF-----VYVIONHQVGNRGK 386
DB 355 ALAKVLURDGFHDGSYSSFRHHGRPI-----NFSAVHPAALVVCSONHQIGNRATG 408

QY 387 ERIKLVDRESYKIAAALYLLSPYIMIPMGEEYGEENPFYFSDFSKLIQGVREGRK 446
DB 409 DRLSOTLPYGSALAAVLTGTGFTPLMLMGEYGASTPWQFFTSHPPELCKATAEGRI 468

QY 447 KE---NGQD---TDPQDESTFNASKLSWKIDEE-----IFSFKILIKMKRELSTAC-- 492
DB 469 KEFERMGWDPVAVPQDPETFRSRKLDWAAEAGDHARLLELYSLTALRRSTPDLTKL 528

QY 493 ---DRRVNVNGENLIIKREVFSLYVFSKSSIEVKVSGTLL 533
DB 529 GFEDTQVAFDEADRWLFRFRGGVQVLLNFSQPVSLDAGATALL 572

RESULT 6
JC4697
trehalose trehalohydrolase (EC 2.4.1.-) - Rhizobium sp.
C:Species: Rhizobium sp.
C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Jun-2000
C:Accession: JC4697
R:Maruta, K.; Hattori, K.; Nakada, T.; Kubota, M.; Sugimoto, T.; Kurimoto, M.
Biosci. Biotechnol. Biochem. 60, 717-720, 1996
A:Title: Cloning and sequencing of trehalose biosynthesis genes from Rhizobium sp. M.
A:Reference number: JC4696; MUID:96219094
A:Accession: JC4697
A:Molecule type: DNA

C:Species: Anabaena sp. (strain PCC 7120)
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AH1827
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, H.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH1827
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-620 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAB77692.1; PID:g17135146; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0168
C:Superfamily: trehalose trehalohydrolase

Query Match 43.3%; Score 1288.5; DB 2; Length 620;
Best Local Similarity 44.0%; Pred. No. 4.2e-76;
Matches 255; Conservative 100; Mismatches 178; Indels 47; Gaps 10;

QY 4 AYKIDGNEVIFTLWAPYQSVKLVL--BKGLYEMERDEKGYFTITLNNVKVRDXYKYL 61
DB 19 AHYLGNGCEGFTWSPSTLNSAVQILKPEKLIPLKAQOAGVQIKVNDVYFGTLRYQL 78

QY 62 DDAEIPDPASRYQPEGVHGSQIIQESKEFNNEF--LKKEDLIYEHVGTFTPEGTF 119
DB 79 NQOEAFAFPASQVOPGEVHGAQVVDHKFEWTDKTSWGSLSLESMIFYELHVGFTFTPEGTF 138

QY 120 EGVIRKLDYKDLGITAEIMPIAQPPGK-----RDWGDGYLYAVONSYGPPG 170
DB 139 TTIIPRLPELRELGINAIELMPIAQPPGDHIEPLAYRNWGYDGYVYPAVQNSYGPAD 198

QY 171 FRKLVDEAHKGLGVLDVYVNHVPEGMVMVKLGPFYSQYKTPWGLTFNFDDEADSEV 230
DB 199 LKNFVNACHENGIAVVLVYVNHFGPEGMVGMQFAPYFTRTYKTPWGNAMFDDAYSQGV 258

QY 231 RKFILENVYIKEYNVGDFRLDAVHAIIIDTSPKHILEEIAADVVKYNR-----RVIAE 284
DB 259 RNYFIONALYWLGEFHIDGLRLDAIQIYDLGAKHFLWELAEAVHNFSGGTWKRHLIAE 318

QY 285 SDLNDPRVVPKCEGYNIDAOQWDDFHHSIHAYLTGRQGYTDFGNLDDIVKSYKDYF 344
DB 319 SDLNNPQIIRPVESGGYGLDAQWSDDFHALLHALTGRQGYQDYDGKCADLAKAYADTF 378

QY 345 VYDGKYNFRKTHGPELVGELGCGNVVYIQNHQVGNRGKERIILKLVDRSYKTAAL 404
DB 379 VYDWRYAPARKRFHGISCDRLPSQFSVCIQNHQDQIGNMQGRLSERISFAGLKLAAGA 438

QY 405 YLLSPYIPMIFMGEYGEENPEYFSDFSKLIQGVREGKKE-----NGQD-TDPQDE 458
DB 439 VLLSPYLPULFMGEYGETAPFIYVSHSDLDLIQAVRAGKEEFPAFYAEDPPDPESA 498

QY 459 STFNASKLSWKIDEE-----IFSFKILIKMRK-----ELSIACDRKRVNVNGE 502
DB 499 EYFLRCKLNLWELHQHVKVLDWYRQLIHLRKTTHPALLNYDRDNEATSDDEKQIVVYR 558

QY 503 NMLIKRGVFSLYVFSKSSIE-----VKYSGTLLLSNNS 538
DB 559 RW--CESREVILAMNPTSPVSLVLTIEKSARKLLDSADS 596

RESULT 5
S65770
maltooligosyltrehalose trehalohydrolase - Arthrobacter sp. (strain Q36)

C:Superfamily: trehalose trehalohydrolase

Query Match 33.5%; Score 997.5; DB 2; Length 580;
Best Local Similarity 38.3%; Pred. No. 3.2e-57;
Matches 212; Conservative 91; Mismatches 215; Indels 35; Gaps 11;
QY 14 FTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVL-DDASEIPDPAS 72
DB 4 FRWAPKPALVRLDV-NGAVHAMRTSADGWHITV-AAPADARYGYLLDDDPVLPDPRS 61
QY 73 RYQEGVHGSPQIIQESKEF---NNET---FLKEDLIIIEIHVGTTPGTEGVRKL 126
DB 62 AROPDGVHARSQSRWEPQPCAAATDTGWPGRSVEGAVIELHIGTGTAGTDAATEKL 121
QY 127 DYLKDLGITAIEIMPIAQFGPKRDMGVDGYLYAVQNSYGGPEGFRKLVDKAEHKKGLGI 186
DB 122 DYLVDLGDIFVLMPPVNSFAGTGWGVDGYLYAVQNSYGGPEGFRKLVDKAEHKKGLGI 181
QY 187 LDVYVNHVGPGEYVNVKLGPFYFQKYPKGLTFNFDDESDVYKFIENVEYWIKEYN 246
DB 182 IDAVFNHLPVSGNVPREFYLSA-SNPWGDGINIAGADSDVHRVHIIIDCALRWMDFH 240
QY 247 VDFRLDAVAHAIIDTSKPHILEEIAADVHVK---RIVAESLNDPRVNVKPKC 300
DB 241 ADGLRLDAVHALVDTTAVHVLLELANATRLSGQLRPLSLIAETDRNDPLITRPSHG 300
QY 301 YNTDAQWVDFHSHIAYLGERGYTDFGNLDDIVKSKYKDFVYDGYKSNFRKTHGE 360
DB 301 YGTAQWDDIHAHITAVSGERGYYADFGSLATLAYTRNGYFHAGTYSFRRRRHR 360
QY 361 PV--GELDGCNFVYIQNHQVGNRGKGERIIKLVDRSYKIAAALYLLSPYIPMFMGE 418
DB 361 ALDTSATPATRELLAYTCTHPOVGNRGLDRPSQYLTGGQAIKALTLGSPYAMLPMGE 420
QY 419 EYGEENPFYFSDSKLIQGVREGKKGQD-----TPQDESTFNASKLSWK-- 469
DB 421 EWGASSPFQFCRPELAHSTVAGRKEEFAEHGWAADDIPDPQDPQRCCKLNWAE 480
QY 470 ---IDEIFSFYKLIKWRKELSTACD-----RRVNVVNGENWLIIGREYFSLYVFSK 521
DB 481 GSGEHARLHRFYDLIALRNEADLPWLDHLMVDYDEQQRVWVMRRGQLMIACNLGAE 540
QY 522 SIEVKYSGTLLLS 534
DB 541 PTCVPVSGELVLA 553

RESULT 9
AI0673
probable hydrolase STY1503 [imported] - Salmonella enterica subsp. enterica serovar Typh
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AI0673
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moulle, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AI0673
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-594 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01762.1; PID:g16502610; GSPDB:GN00176
C:Genetics:
A:Gene: STY1503
C:Superfamily: trehalose trehalohydrolase

Query Match 32.9%; Score 978.5; DB 2; Length 594;

Best Local Similarity 43.3%; Pred. No. 5.7e-56;
Matches 209; Conservative 76; Mismatches 169; Indels 29; Gaps 12;
QY 7 IDGNEVITLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLLDDASE 66
DB 15 IADVVFRFLWATGACQOQVLMRLAGKD-OEMOASGDGFTLDVSGVTPGTETNFVLSGMY 73
QY 67 IPDPASRYQEGVHGSPQIIQESKEFNNETFLKK---EDLIIIEIHVGTTPGTEGFGVI 123
DB 74 LPDPASRAQKTDVNGPSYVIDPGSYAWRNTGWKSRWEQAVVYEMHTGTTPGTFHTAI 133
QY 124 RKDYLKDLGITAIEIMPIAQFGPKRDMGVDGYLYAVQNSYGGPEGFRKLVDKAEHKKGL 183
DB 134 AKLPYLAELGVTVIEVMPVAQFGGERGWYDGYLLYAPHASAYGTPDDKAFIDTAHGYGL 193
QY 184 GVILDDVYVNHVGPGEYVNVKLGPFYFQKYPKGLTFNFDDESDVYKFIENVEYWI 242
DB 194 SVILDDVYVNHVGPGEYVNVKLGPFYFQKYPKGLTFNFDDESDVYKFIENVEYWI 250
QY 243 KEYNVDCGFRLDVAHAIIDTSKPHILEEIAADVHVK---YNRIVIAESLNDPRVNVKPKC 299
DB 251 TEYHLDGLRDAIDQIEDSSARHVLVEIAQRIREDITDRPIHLTTEDSRNIISLHPRDQD 310
QY 300 GYN--IDAQWVDFHSHIAYLGERGYTDFGNLDD--IVKSYKDFVYDGYKSNFR 355
DB 311 GNAPLFTAENDDDFHNAVHVAFATGETQAYYNDFADEKPEKHLARALAEQFAYQGEIS---- 366
QY 356 KTHCEPVGELD---GCNFVYVYQNHQVGNRGKGERIIKLVDRSYKIAAALYLLSPYI 411
DB 367 POTGEPGVKSTGPPVAFVDFIQNHQVGNRAQGRLLITLAGAERTKVLATLLLSPHI 426
QY 412 PMTFMGEEYGEENPFYFSDSKLIQGVREGKKGK---NGQDT-DPQDESTFNASKL 466
DB 427 PLFMGEYGESRPFLETFDF-HCDLARAVREGRAKEFADHAGENVDPNAPETFORSKL 485
QY 467 SWK 469
DB 486 NWK 488

RESULT 10
H75516
probable maltotriose trehalose trehalohydrolase - Deinococcus radiodurans (strain R1
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: H75516
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: H75516
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-600 <WHI>
A:Cross-references: GB:AE001905; GB:AE000513; NID:g6458144; PIDN:AAF10042.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0464
A:Map position: 1
C:Superfamily: trehalose trehalohydrolase
Query Match 30.4%; Score 905.5; DB 2; Length 600;
Best Local Similarity 38.7%; Pred. No. 3.2e-51;
Matches 205; Conservative 86; Mismatches 170; Indels 69; Gaps 20;
QY 14 FTLWAPYQKSVKLVK-----LEKGLYEMERDEKGYFTITLNNVKVRDRYKYVLLDDA 64
DB 38 FRLTWTARTAVVRVNGTEHVMSTLGGGIYELELP-----VGPGRYLFVLDGV 86
QY 65 SEIPDPASRYQEGVHGSPQIIQESKEFNNETF-----LKKEDLIIIEIHVGTFTTP 115

Db 87 -PTDPYARFLPDGVHGAEEVY-----DFG--TFDWTADWHGIKLADCVFVEVHVGTTTP 139
Qy 116 EGTFGVIRKLDYKLDIGTAEIMPIAQFPGKRDWGDVYLVAVQNSYGGPEGRKLV 175
Db 140 ECTYPAAEKLPYLKELGVTALQVMPLAFAFGQGWGDGAIFYAPAYGRPELMAVL 199
Qy 176 DEAHKKGLGVLDVYVNVHVGPEGNYMVKLGP-YFSQKYKTPWGLTFNFDAAESDEVRKFI 234
Db 200 DAAHRLGLGVLDVYVNVHFGPSGNSLVAPSYFTDRESSAWGM--GLDYAE-PMHRRYV 256
Qy 235 LENVEYWKENVGDGFRLDVAHAIIDTSPKHILEBIAIDVAVKQY--NRVIVAESLNDNDPRV 292
Db 257 TGNARWLRLDYHFDGLRLDAYPTMTDDSETHILTAQEIHELGGTTHLLAEHDHRLNPLD 316
Qy 293 VNPKEKCCYNIDAQWDDFHSHIAYLGRQGYTYTDF-GNLDDIVKSYKDVVYVDGKYS 351
Db 317 VTVN-----HLDGIWDDFHETRTVLTGEGQYVAGYRGGAELAYTRGWRYYEGQFW 371
Qy 352 NFRKTH--GEPVGLDGCNFVYIQNHQVGNRGKGERIILK--VDRESYKIAAALYL 407
Db 372 AVKGEHERGHPSDALEAPNFVYIQNHQVGNRGKGERIILK--VDRESYKIAAALYL 430
Qy 408 SPYIPMIFMGEGYGEENPFYFSDSKLIQGVREGKKKE-----NGQDT-DPQDES 459
Db 431 -PMTPLLFQGOEWAASSTPFQFSDHA-GELGQAVSEGRKKBFSGFSGEDVPDPQAEQ 488
Qy 460 TFNASKLSWKIDE-----EIFSFKYKILMKRKELSIACDR-RVNVVNGEN 503
Db 489 TFLNSKLNWAEREGEHARTLRLYNDLRLREDPVLHNRQRENTTGH 538

RESULT 11
G83375
probable glycosyl hydrolase PA2164 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83375
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337
A:Accession: G83375
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-583 <STO>
A:Cross-references: GB:AF004643; GB:AE004091; NID:g9948178; PIDN:AG05552.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2164
C:Superfamily: trehalose trehalohydrolase

Query Match 29.6%; Score 881.5; DB 2; Length 583;
Best Local Similarity 36.3%; Pred. No. 1.1e-49;
Matches 203; Conservative 95; Mismatches 174; Indels 87; Gaps 19;

Qy 3 FAYKIDGN-EVIFLWAPYQKSVKLKLEKLYEMERDEKGYFTITLNNVKVRDRYKVL 61
Db 7 FGAQFOGNGRTCFGWLWAPDAREVRVETADGRDWPLEGSDSGWFATL-PCPPGTRYRYRI 65
Qy 62 DDAEIPDPASRYOPEGVHGHPSQIQES-----KEFNNETFLKKDLIYIEIHVGFTTP 115
Db 66 DGRGVDPASQFQPDGVHGHSHQVLDHGTAYARVDEWRGRW---HEAVIELHVGFLF-- 120
Qy 116 EGTFGVIRKLDYKLDIGTAEIMPIAQFPGKRDWGDVYLVAVQNSYGGPEGRKLV 175
Db 121 -GSAVERFRLPRVELGVTAVELMPLGEFPGRRWGVDGLVLPAPASAYCTPEQLKHLI 179
Qy 176 DEAHKKGLGVLDVYVNVHVGPEGNYMVK-LGPFYSQYKTPWGLTFNFDAAESDEVRKFI 234

Db 180 DSAHGMGLMVFVDVIYNHFGPDGNYLAQYAAAFRRDQTPWQQAIDF---RRGEVREFF 236
Qy 235 LENVEYWKENVGDGFRLDVAHAIIDTSPKHILEBIA-----DWHKYNRIVIAE 284
Db 237 YENALMWLLDVRDGLRFDVAHAIPDSA---FLVEMARLRCAAGPERHVH-----LVLE 288
Qy 285 SDLNDPRVYVNPKEKCCYNIDAQWDDFHSHIAYLGRQGYTYTDFGN-LDDIVKSYKDV 343
Db 289 NDDNRASLL---RQGY--DAQWDDGHHLVLLTGENDGYQDYPEPLRLARCLAE 342
Qy 344 FVYDGKYSNFRKTHGEPVGLDGCNFVYIQNHQVGNRGKGERIILKVDRESYKIAA 403
Db 343 FVYQGE-ANRHRGRPEESADLPADFVLFQNHQVGNRAFGKLSVLAEPQALRLAIA 401
Qy 404 LVLSPYIPMIFMGEGYGEENPFYFSDSKLIQGVREGKKENGOD----- 452
Db 402 LQLAPMPLPLFMGECAAREFFLYFTD-HOGELADAVREGRRKEFGFGRFGEGATLAS 460
Qy 453 -TDQDESTFNASK-----LSWKIDEEIFSYKILMKRKELSI----- 490
Db 461 LPDPNAVETFEERSRPLADCDPAWR-----GFYROLLEIRHEHLIPRLRGARSLGVTTI 514
Qy 491 ---ACDRRVNVVNGENWLI 506
Db 515 AGAALSARWRLGDSGSDWRI 533

RESULT 12
D96001
probable 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) [imported] - Sinorhizobium m
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
C:Accession: D96001
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmaster, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: D96001
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIDN:CAC49676.1; PID:g15141163; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
C:Contents: annotation
C:Genetics:
A:Gene: gigB2; Smb21447
A:Genome: plasmid
C:Superfamily: trehalose trehalohydrolase
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 29.4%; Score 876.5; DB 2; Length 601;
Best Local Similarity 35.3%; Pred. No. 2.5e-49;
Matches 207; Conservative 99; Mismatches 204; Indels 77; Gaps 19;

Qy 7 IDGNEVFTLWAPYQKSVKLKLEKLYEMERDEKGYFTITLNNVKVRDRYKVLDDASE 66
Db 19 IDSTCFRLWADEREVDL-VLGGAVHKMQSLDGGWFEITL-AKAGERICFRADGTE 76
Qy 67 IPDPASRYQPEGVHGHPSQIQESK-EFNNETFLKK--EDLIYETHVGTFTPEGTFGVI 123
Db 77 VADPASSAQEREASGTSIVVDHAAVEMQTSWRGRPWEEAVISELHVGCFTPEGFFRAAI 136
Qy 124 RKLKLDLGLTGITATEIMPIAQFPGKRDWGDVYLVAVQNSYGGPEGRKLVDEAHKKGL 183

Db 137 ERLPHLAGAGITAIEIMPVQAQFVGRGWGVDGLVHYAPHNAYGKPPDDLKALVDAASHLGL 196
QY 184 GVTLDVVYVNHVGGPMVYKLG-P-YFSOKYKTPWGLTFNFDDAESDEVKRFILNVEYWI 242
Db 197 TVLLDDVVYVNHVGGPMVYKLG-P-YFSOKYKTPWGLTFNFDDAESDEVKRFILNVEYWI 253
QY 243 KEYNDVGFRLDAVHAIDTSPKHILEIADVVHK---YNRIVIAESDLNDPRVVPKPKC 299
Db 254 GHFRFGLRLDATEIQIRDTTKPHFLVALEHVEFAERQIHLVLEDAHRRSLQORDAS 313
QY 300 GYNI--DAQWDDFHSHIAYLTGERQGYTDFGN--LDDIVKSYKDVYVYDGYKYSNFR- 354
Db 314 GARMLEDAWDDHNLHNVATGETKGYRLFADEPWGKIRSALEAGFAVPAKEDNFSF 373
QY 355 --RKTGCEPVGELDGCNFVYIQNHQDVGNRGKGERIILVDRESYKIAAALVLLSPYIP 412
Db 374 EGSRAVPQGRVN-----FLQNHQDIGNRAFERGLASLQEDSLRVLAAMHMLTPQIP 427
QY 413 MIFMGEYGEENPFYFSDSFKLIQVREGKKE-----NQGD-----TDQDEST 460
Db 428 LLFMGEYGETOPFYFSDY-QGEIAAAIRLGRDEAENFGGLPECKTMDLDPDLPDPV 486
QY 461 FNASKLSWK-----IDEEFSPFKILMKR----- 486
Db 487 FAGSKLRNRATSPAGERHLAVYRDIAVIRORHIVPLIAGTAVPDGRAYETKDGIIAVDW 546
QY 487 ELSIAC-DRRVNVNNGENWL-IKRGYFSLYVFSKSIEVKYSGTL 531
Db 547 QFGEACLELRVNLHETHAVPAIRGQPIFT---SETSSVETVYVYSGEL 590

RESULT 13
AG3194
alpha-amylase Atu5284 [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plasmid
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C:Accession: AG3194
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AG3194
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-595 <KUR>
A:Cross-references: GB:AE008687; PIDN:AAL45973.1; PID:g17743726; GSPDB:GN00188
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu5284
A:Genome: plasmid
C:Superfamily: trehalose trehalohydrolase

Query Match 28.9%; Score 860.5; DB 2; Length 595;
Best Local Similarity 34.7%; Pred. No. 2.7e-48;
Matches 206; Conservative 89; Mismatches 246; Indels 53; Gaps 16;
QY 4 AYKIDGNEVIFTLWAPYQKSVKLVLEKGLYMERDEKGYFTITLNNVVRORYKYVLDD 63
Db 12 AQHQSGDFFSIWAPASATVKL-WLNDAEFDMHTAGDGHWDIT-KPALPGDRYGVFLAD 69
QY 64 ASEIPDPASRYQPEGVHGVSQIIQESKEFNFTLKK--EDLIYBIHVGTFTEPTFEG 121
Db 70 RTRVADPASRQOEGPRGFSLLVNHDFANKPNWKGPRWHEAVVYELHIGTFTEPTFAA 129
QY 122 VIRKLDYLKDLGITAIEIMPIAQFPGKRDWGDVYLVAVQNSYGGPEGKRLVDEAHKK 181
Db 130 AAKELEYLADVGTITIELPLATFAGSRGWGDVGLQFSPQDYGTPDELKAFIDQAHGH 189

QY 182 GLGVILDVVYVNHVGGPMVYKLG-P-YFSOKYKTPWGLTFNFDDAESDEVKRFILNVEYWI 241
Db 190 GIMVLLDVYVNHVGGPMVYKLG-P-YFSOKYKTPWGLTFNFDDAESDEVKRFILNVEYWI 246
QY 242 IKENYVDGFRDLDAVHAIDTSPKHILEIADVVHK---EIAADVHKYNRIVIAESDLNDPRVVPK 296
Db 247 LETYREDGLRIDAADAHLAGDGEVDFLIEMAREVKRTIRNRHVHLVIEDARNAASPTM 306
QY 297 ERGCGYNDAQWDDFHSHIAYLTGERQGYTDFGN--LDDIVKSYKDVYVYDGYKYSNFR 354
Db 307 ADGAILVDAQWDDFHSHIAYLTGERQGYTDFGN--LDDIVKSYKDVYVYDGYKYSNFR 366
QY 355 R-KTHGCEPVGELDGCNFVYIQNHQDVGNRGKGERIILVDRESYKIAAALVLLSPYIP 413
Db 367 NFAASGEPGSHLPFHRFVFNHNDQAGNRLRGERLALIPPLFCTLEAILLLCQPTPL 426
QY 414 IFMGEYGEENPFYFSDSFKLIQVREGKKE-----NQGD-----TDQDESTN 462
Db 427 VFMGDEHGSANPFFFSDFHPDHNREQEIRNLKQAESFQCELPDASQMVMDPDQHTMQ 486
QY 463 ASKLSWKIDEEI-----FSFYKILMKR-----KELSIACDRRVNVNNGENW 504
Db 487 LSTLKWTHAETEGROSADMAALLAKRRGHVPLLCSHFEKGISLDCCEPRCLAI--DW 543
QY 505 LIITGREYFSLYVFSKSIEVKYSGTLSSNNSPFQHIIEEGKYE-FDKGFALY 557
Db 544 HFKAGRLMRANL-SENMCCLPAVKGEILLRNGS-----VGNTRYEGYAAQFAIY 592

RESULT 14
S19134
probable alpha-glucanotransferase - Anabaena variabilis
C:Species: Anabaena variabilis
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
C:Accession: S19134
R:Sato, N.
Plant Mol. Biol. 18, 165-170, 1992
A:Title: Cloning of a low-temperature-induced gene lti2 from the cyanobacterium Anaba
A:Reference number: S19133; MUID:92119230
A:Accession: S19134
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-552 <SA>
A:Cross-references: EMBL:X59279; NID:g39252; PIDN:CAA41970.1; PID:g39253
C:Superfamily: neopullulanase, alpha-amylase core homology

Query Match 17.6%; Score 524.5; DB 2; Length 552;
Best Local Similarity 28.0%; Pred. No. 1.7e-26;
Matches 148; Conservative 90; Mismatches 162; Indels 129; Gaps 22;
QY 14 FTLWAPYQKSVKLVLEKGLYE--MERDEKGYFTITLNNVVRD---RYKYVLDDAS--- 65
Db 7 FTLFAPYKGAALIASFSQWQEIIPMKDGDGYFRFT---VELEDGTYQKFRVQTRSWFF 63
QY 66 -----EIPDPASRYQPEGVHGVSQIIQESKEFN--ETFLKED-----LIYBI 108
Db 64 EEDQWVDVTDPTATDIDES-SGKDNSIARIKDKGEKIVYVWQHDQKPLPADHELVIYEL 122
QY 109 HVGTFTE-----PECTFEGVIRKLDYLKDLGITAIEIMPIAQFPGKRDWGDVYLVAV 161
Db 123 HVGDFSGGDDDPYARGKYKHVIEKLDYLCGLGINALTELLPVKEYPGDYSWGNPRYFFAT 182
QY 162 QNSYGGPEGKRLVDEAHKKGLGVILDVYVNVHVPBGNY-MVKLGYFYSQKYKTP---WG 217
Db 183 ESSYGTADLKKLVLDCHORGIIRIMDGIYNHSEASSPLTQIDHDYVHHEPRDPDNNNG 242
QY 218 LTFNE---DDAESDEVKRFILNVEYWKYKYNVDGFRDLDAVHAIDTSPKHL---EE 269
Db 243 PEFNYEHYDENLETYPARKFIGDTRVRYWVGVEYHLDGIRYDAARQIANYDFMHWIAQEA 302
QY 270 IADVHVHKYNRIVIAESDLNDPRVNV---PKKCGYNIADAQWVDDFHHSTHAYLTGRQCY 326

Db 303 TAGAKPFYN---VAEHIPETTSITNLDGPMDCG-----WHDSPYHTIKAHICGDT--- 349
Qy 327 YTFGNLDDIVKSKYKDVYVYDGKYSNFRKTHGEPVGELDCNFFVYIQNH-----QV 380
Db 350 -FDLENLKDVI-----DPKRQGF-----LGATNVVNYLTNHDHIMVEL 388
Qy 381 GNRGGERIILKLVDRSY---KIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDFSLSKL 437
Db 389 GNR-----EIFHDEAFRAKLGTAIIMTAVGVPLIWMGEFGEYKP----- 429
Qy 438 IQGVREGRRKKGQDTPDQDSTFNASKLSWKIDEEIFSYFKILIKMRK 486
Db 430 -----KQDQSKIDWTLLGNDLNRSLFDYHKGLIGLRK 462

RESULT 15

AH1915
hypothetical protein all10875 [imported] - Anabaena sp. (strain PCC 7120)
C:Species: Anabaena sp.
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AH1915
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH1915
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-552 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAE72832.1; PID:gl7130220; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all10875
C:Superfamily: neopullulanase; alpha-amylase core homology

Query Match 17.6%; Score 524.5; DB 2; Length 552;
Best Local Similarity 28.0%; Pred. No. 1.7e-26;
Matches 148; Conservative 90; Mismatches 162; Indels 129; Gaps 22;

Qy 14 FTLNAPYQKSVKLVLEKGLYE--WERDEKGYFTITLNNVKVRD---RYKYVLDDAS--- 65
Db 7 FTLFAPYKGAALIASFSDMWEIPMKKGGDGYFRFT---VELEDGTQYKFRVQTRSWFF 63
Qy 66 -----EIPDPASRYQPEGVHGPSQIIQESKEFNN--ETFLKKED-----LIIVEI 108
Db 64 EEDQWVDVTDYATDIDES--SGKDNSTARIKGEKIVDTYVQWQHDKPLPADHELVIYEL 122
Qy 109 HVGTFT-----PEGTGFEVIRKLDYKDLGITAIEIMPIAQPPGKRDWGDGVLYIYAV 161
Db 123 HVGDFSGGEDDPYARGVKYKVKIEKLDYLCELGINAIELLPVKYKPGDYSWGNRYFFAT 182
Qy 162 QNSYGGPEGFRKLVDEAHKGLGVILDVNVNHHVGPENY--MVKLGPYFSOKYKTP---WG 217
Db 183 ESSYGSTADLKKLVDECHQRGIRIIMDGIYNHSEASSPLAQIDHDYWHHEPRDPDNNWG 242
Qy 218 LTFNF---DDAESDEVKRFLENVEYKWEYNVDGFRDLDAVHAIDTSPKHIL-----EE 269
Db 243 PEFNVEHYDENLEYTPARKFIGDVTYVWGEYHLDGIRYDAAQIOANYDFMHWIAQEAKK 302
Qy 270 IADVVKYKIRIVIAESDLNDRPVVN---PKEKGCYNIDAQWDDFHHSIHAYLTGERQGY 326
Db 303 TAGAKPFYN---VAEHIPETTSITNLDGPMDCG-----WHDSPYHTIKAHICGDT--- 349
Qy 327 YTFGNLDDIVKSKYKDVYVYDGKYSNFRKTHGEPVGELDCNFFVYIQNH-----QV 380
Db 350 -FDLENLKDVI-----DPKRQGF-----LGATNVVNYLTNHDHIMVEL 388
Qy 381 GNRGGERIILKLVDRSY---KIAAALYLLSPYIPMIFMGEEYGEENPFYFFSDFSLSKL 437
Db 389 GNR-----EIFHDEAFRAKLGTAIIMTAVGVPLIWMGEFGEYKP----- 429

Qy 438 IQGVREGRRKKGQDTPDQDSTFNASKLSWKIDEEIFSYFKILIKMRK 486
Db 430 -----KQDQSKIDWTLLGNDLNRSLFDYHKGLIGLRK 462

Search completed: July 15, 2002, 12:18:43
Job time: 135 sec


```
Db 59 APTDGNVD-----YGLLDGDETPLDPTRRQPGDGVHLSRTFDPFSAYSWQDDAQWQREL 114
Qy 102 -DLIIYEHVGFTEPGEFVIRKLDYLDKLGITAITAEMPFAOPGKRDMGVDGYLYA 160
Db 115 QGAVIYELHGTFTPEGTEAAAGKLDYLAGLVDFIELLPVNAFNGTHNGYDGVQWFA 174
Qy 161 VQNSYGGPGEFKLVDEAHKGLGVLVDVYVNHVGPENYMKLGPYFSQKYKTPWGLTF 220
Db 175 VHEAYGPEAYORFVDAHAAGLVQDVVYVNHVHLGPGSNYLPREFGPKLQGEQNTWGDV 234
Qy 221 NFDAESDEVKRFLENVVEYIKENYVDGFRDVAHALIDTSPKHILBEIADVVHKYNI 280
Db 235 NLDGPGSDHVRVYILDNLAMWLRDYRVDDLDAVHALKDERAVHILEDFGALADQISAE 294
Qy 281 V-----IAESDLNDPRVVPKCKGYNIDAQWDDFHSHIAYLTGERQCYTYTDFGNLD 334
Db 295 VGRPLTLIAESDLNPNRLLYPRDVNGVLEQWSDDFHHAHVNVVGTGTYTDFDLSA 354
Qy 335 DIVSKYKDVYDGYKSNFRKTKHGPVGLDGCNF-----VYIQNHQDVGNRKGK 386
Db 355 ALAKVLDRGFHDGYSFSFRHHRPI-----NFSAVHPAALVVCSONHDQIGNRATG 408
Qy 387 ERIIKLVDRSEYKTAALYLLSPYIPMFWMGEYEEENPFYFSDFSKLIQVREGRK 446
Db 409 DRLSQTLPYGLSALAVALTLTGPTPTMLLMGEYEGASTPWQFFTSHPPELGKATAEGRI 468
Qy 447 KE-----TDPODESTFNASKLSWKIDEE-----IFSEYKILIKMKKELSIAC-- 492
Db 469 KEFERMGWDPAVVPDPQDPETFRSKLDWAEAGDHARLLELYRSLTALRSTPDLTKL 528
Qy 493 ---DRRVNVNGENWLIKIGREYFSLYVFSKSSIEVYKSGTLL 533
Db 529 GFEDTQVAFEDARWLRFRRGGVQVLLNFSEQPYSLDGAGTALL 572

RESULT 2
TREZ_RHISP STANDARD; PRT; 596 AA.
AC Q53238;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Malto-oligosyl trehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE alpha-D-((1->4)-alpha-D-glucano)trehalose trehalohydrolase)
DE (Maltooligosyl trehalose trehalohydrolase).
GN TREZ.
OS Rhizobium sp. (strain M-11).
OC Bacteria; Proteobacteria; alpha subphylum; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=391;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96219094; PubMed=8829547;
RA Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
RT "Cloning and sequencing of trehalose biosynthesis genes from Rhizobium
RT sp. M-11.";
RL Biosci. Biotechnol. Biochem. 60:717-720(1996).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC linkage in 4-alpha-D-((1->4)-alpha-D-glucanosyl)(n) trehalose to
CC yield trehalose and alpha-(1->4)-D-glucan.
CC -1- PATHWAY: Trehalose biosynthesis.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; D78001; BAAL1187.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase.
FT ACT_SITE 265 265 BY SIMILARITY.
SQ SEQUENCE 596 AA; 65262 MW; 73EA90AE0534DDCD CRC64;

Query Match 34.4% Score 1025; DB 1; Length 596;
Best Local Similarity 40.3%; Pred No. 8.4e-60;
Matches 223; Conservative 80; Mismatches 197; Indels 62; Gaps 16;

Qy 14 FTLWAPYQKSKYKLVLEKG-LYEMER-----DEKGYFTITLNNVKVRDRYKYVLD-DAS 65
Db 17 FDIWAPAGTTL--LAGGERYEMRRPGNGPADGWTAAADAPTGDADVDYGYLLDGEI 74
Qy 66 EIPPASRYQPEGVHGPSQIIQESKEFNNEFLAKEDL---IYIEIHVGFTEPGEV 122
Db 75 PLDPRTTRRQPEGVHLSRTFDPGAHRWQDAGWQRELGQSVIYELHGTFTPEGLDAA 134
Qy 123 IRKLDYLDLIGITAIEIMPFAOPGKRDMGVDGYLYAVQNSYGGPGEFKLVDEAHKKG 182
Db 135 AGKLDYLAGLIDFIELLPVNAFNGTHNGYDGVQWFAVHEGYGPAAYQRFVDAHAAG 194
Qy 183 LGVILDVYVNHVGPENYMKLGPYFSQKYKTPWGLTFNFDDAESDEVKRFLENVYWI 242
Db 195 LGVIQDVVYVNHVGPENYMKLGPYFSQKYKTPWGLTFNFDDAESDEVKRFLENVYWI 254
Qy 243 KEYNVDFRLDVAHALIDTSPKHILEE---IADVVKYN---RIVIAESDLNDRVVPNK 296
Db 255 RYRVGDLRLDAVHALKDERAVHILEEFGALADALSSEGGRLPLTIAESDLNPNRLLYPR 314
Qy 297 EKCYNIDAQWDDFHSHIAYLTGERQCYTYTDFGNLDIVKSKYKDVYVVDGKYSNFRK 356
Db 315 DVNGYGLAQWSDDFHHAHVNVVSGETGYTSDFSLGALAKVLDRDFFHDSYSSRGR 374
Qy 357 THGPEVGLDGCNF-----VYIQNHQDVGNRKGRIIKLVDRSEYKIAAALYLLS 408
Db 375 CHGRPI-----NFSAVHPAALVVCSONHDQIGNRATGDRLSQSLPYGLSALAVALTLTG 428
Qy 409 PYIPMFWMGEYEEENPFYFSDFSKLIQVREGRKE-----NGQD-----TDPODESTF 461
Db 429 PFTPLMFWMGEYGAATPWQFFTSHPPELGKATAEGRIREFRMGWDPAVVPDPQDPETF 488
Qy 462 NASKLSNK-----IDEEIFSEYKILIKMKR-----ELSIACDRRVNVNGENWLIKGR 510
Db 489 TRSKLDWAEASAGDHARLLELYRSLTLRSTPELARLGEA-DTAVEFDDDAEWL----- 542
Qy 511 EYFS-----LYVFSKSSIEVYKSGTLL 533
Db 543 RYMRGGVQVVLNFADRPISLDRPGTALL 570

RESULT 3
TREZ_ARTRM STANDARD; PRT; 575 AA.
AC Q9AJN6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Malto-oligosyl trehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE alpha-D-((1->4)-alpha-D-glucano)trehalose trehalohydrolase)
DE (Maltooligosyl trehalose trehalohydrolase).
GN TREZ.
OS Arthrobacter ramosus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcales; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=1672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S34;
RA Yamamoto T., Maruta K., Watanabe H., Yamashita H., Kubota M.,
RA Fukuda S., Kurimoto M.;
```

```
RT "Trehalose producing operon treYz from Arthrobacter ramosus S34." ;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to
CC yield trehalose and alpha-(1->4)-D-glucan.
CC -!- PATHWAY: trehalose biosynthesis.
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB045141; BAB40766.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase.
FT ACT_SITE 250 250
SQ SEQUENCE 575 AA; 63079 MW; 284A3F20207E228B CRC64;

Query Match 34.4%; Score 1024.5; DB 1; Length 575;
Best Local Similarity 39.4%; Pred. No. 8.6e-60;
Matches 215; Conservative 90; Mismatches 189; Indels 51; Gaps 13;

QY 14 FTLWAPYQSKVLKLEKGLYE--MERDEKGYFTITL---NNVYRDRYKYYLDDASEIP 68
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
5 FFWAPQAAQVTL-VVGQRAELPLTRDENGWALQPPWDGGGPDLDV-YGYLVDGKGPEA 62
QY 69 DPASRYQPEGVCHGPSIIQESKEFNETFLKED-----LIYEIHVGTFTPEGTF 119
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
63 DPLSLRQPPGVH-----ELGRFDPARYAGDGDGWRGDLTGAVIYELHVGTFTEGTL 116
QY 120 EGVIRKLDYKIDGITAIETIMPTAQPPGKRDWGYDGYLYAVONSYGGPEGFKLVDEAH 179
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 DSAIRLHLVRLGVDVAVELLPVNAFNGHWGVDGLVAVHVPKGGPEAQRFVDACH 176
QY 180 KGLGVILDVYVNHVGPENYVYKLPYFSQYKTPWGLTFNFPDDAESDEVKRFILNVE 239
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 ARGLAVVQDVYVNHVGPENYVYKLPYFSQYKTPWGLTFNFPDDAESDEVKRFILNVE 236
QY 240 YMIKEYNVDFRLDAVHAIDTSPKHLEIA----DVVHKYR--IVTAESDLNDPRVY 293
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
237 YLWRDMHADGLRLDVAHLRDLARALHLEELAAARVDDELAGELGRLPTLIAESDLNDPKLI 296
QY 294 NPKEKCGYNIDAQWDDFHSHIAYLTGRQGYTDFGNLDDITVSKDVFYVDGKYSNF 353
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
297 RSRAAHGYGLDAQWDDVHVAHVHANYGTETVGYADFGGLGALVKYVQRGWFDGTWSF 356
QY 354 RKKTGHEPVG-ELDGCNFVYVYQNHQDVGNRGKGERIKLVDRSYKIAAALYLSPYTP 412
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
357 RERHHGRPLDPTFRRLLVAFADHDQVGNRAVGDRAVSAQVGGEGSLAAALVLLGPFTP 416
QY 413 MLFPMGEYGEENPFYFSDSKLQGVREGRKKE---NGQD----TPQDESFYNASK 465
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
417 MLFPMGEWGARTPQWQFTTSHPPELGEATARGIAEFARWGMDPAVVPDPQATFARSH 476
QY 466 LSWKIDEE-----IFSFKYLIKMKRKLSTACIDRNVVNVNGENWLLIKGREYFSLVFSK 520
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
477 LQNSEPERPHAGLAFYDTLALRELVPDAPARE-----VDADAEARGVFAFSR 526
QY 521 SSIEV 525
Db 1 :
527 GPLRV 531

RESULT 4
TREZ_MYCTU
ID TREZ_MYCTU STANDARD; PRT; 580 AA.
```

```
AC Q10769;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE alpha-D-[(1->4)-alpha-D-glucano]trehalose trehalohydrolase)
DE (Maltooligosyl trehalose trehalohydrolase).
GN TREZ OR RV1562C OR MT1613 OR MTCY48.03.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to
CC yield trehalose and alpha-(1->4)-D-glucan.
CC -!- PATHWAY: Trehalose biosynthesis.
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z74020; CAA98329.1; -.
DR EMBL; AE007027; AAK45880.1; ALT_INIT.
DR TIGR; MT1613; -.
DR Tuberculist; Rv1562c; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Hydrolase; Glycosidase; Complete proteome.
FT ACT_SITE 247 247
FT ACT_SITE 247 247 BY SIMILARITY.
SQ SEQUENCE 580 AA; 64076 MW; 506190468F4B862 CRC64;

Query Match 33.5%; Score 997.5; DB 1; Length 580;
Best Local Similarity 38.3%; Pred. No. 5.1e-58;
Matches 212; Conservative 91; Mismatches 215; Indels 35; Gaps 11;

QY 14 FTLWAPYQSKVLKLEKGLYEMERDEKGYFTITLNNVYRDRYKYYL-DDASEIPDPAS 72
Db 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 FRWAPKPALVRLDV-NGAVHAMTSADGWHTTV-AAPADARYGYLLDDDDPTVLDPKRS 61
```

```
QY 73 RYQEGVHGSPQIIQESKEF---NNET---FLKKEDLIIVHIGVTFTEGFEVIRKL 126
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 62 ARQPDGVHARSORWEPQCFGAARTDTGWPCRSVEGAVIELHIGTFTTAGTFDAAEKL 121
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 127 DYLDLGLTATEIMPIAQFPGKRDWGDVYLYAVQNSYSGSGEPGFKLVDEAHKGLGVI 186
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 122 DYLDLGLDIFDELMPVNSFAGTRGMGVDGLVWYSHVEPYGPGDGLFRFIDACHARRLGLV 181
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 187 LDVYNNHVGPGNVMWKLGPYFSQKYKTPMGLTFNFDDAESDEVKRFLEINVEYWIKEYN 246
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 182 IDAVFNHLGSGNVLPRFGPYLSSA-SNPWGDGINIAGADSEVRYIIDCALRMRDFFH 240
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 247 VDGRLDAVHAIDTSPKHILEETADVVHKNY-----RIVIAESLNDLPRVVPKPKCG 300
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 241 ADGLRLDAVHALVDTTAVHVEELANATRWLSGGLRPLSLIAETDRNDRLITRPSHGG 300
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 301 YNIDAQWVDFHHISIHAYLTCERGQYTDGRLNDLDDIVKSYKDVYVQNSYSGSGEPGFKLV 360
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 301 YGITAQWDDIHAHTAVSGERQGYADFGSLATLAYTLRNGYFHAGTYSSFRRRHR 360
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 361 PV--GELDGCNFVYIQNHQVGNRGKRIIKLVDRSYKIAAALYLLSPYIPMIEMGE 418
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 361 ALDTSATPATRLAYTCTHQVGNRGLDRPSQVLTGGLAKAALYLLSGPYTAMLENGE 420
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 419 EYGEENPFYFSDSFKLIQVREGKENGQD-----TDPQDESTFNASKLSWK-- 469
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 421 EWGASSPFQFCSHPELAHSTVAGRKEEFAEHGWAADDIPDPQDPTQRCCKLNWAAE 480
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 470 ---IDEIFSFYKILIKMRKELSTACD-----RRVNVVNGENWLLIKGREYFSYVFSKS 521
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 481 GSGEHALHFRFYRLIALRNEADLADPLDHLMDYDEQQRVWYMRGGLMIACNLGAE 540
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 522 SIEVYSGTLLS 534
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 541 PTCVPVSGELVLA 553
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:

RESULT 5
TREZ_BREHE STANDARD; PRT; 589 AA.
AC 052520;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE alpha-D-(1->4)-alpha-D-glucanotrehalose trehalohydrolase)
DE (Maltooligosyl trehalose trehalohydrolase).
GN TREZ.
OS Brevibacterium helvolum.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcales; Brevibacteriaceae; Brevibacterium.
OX NCBI_TaxID=1704;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11822;
RA Kwon T., Kim C.H., Choi Y.D.;
RT "Maltooligosyl trehalose synthase/trehalohydrolase encode proteins for
RL trehalose production in Brevibacterium helvolum";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC linkage in 4-alpha-D-[(1->4)-alpha-D-glucanosyl](n) trehalose to
CC yield trehalose and alpha-(1->4)-D-glucan.
CC -|- PATHWAY: Trehalose biosynthesis.
CC -|- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF039919; AAB95369.1; -.
DR InterPro; IPR000461; Alpha_amlase.
DR Pfam; PF00128; alpha-amylose; 1.
KW Hydrolase; Glycosidase.
FT ACT_SITE 258 258 BY SIMILARITY.
SQ SEQUENCE 589 AA; 64217 MW; C282314B6E9BA029 CRC64;

Query Match 32.4%; Score 965; DB 1; Length 589;
Best Local Similarity 38.5%; Pred. No. 6.9e-56;
Matches 218; Conservative 88; Mismatches 216; Indels 44; Gaps 13;

QY 14 FTLWAPYQKSVKLVLEKGLYMERDE-----KGYFTITLNNVVRDRYKYVLD-DASEI 67
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 11 FDWAPDVSSVL-VADGRQYPMQKKEAPGSEGWWTASDAPPNGVDYDGYLLDGNTPV 69
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 68 PDPASRYQPEGVHGSPQII-----QESKEFNNETFLKKEDLIIVHIGVTFTEGFE 120
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 70 PEPSSRLPAGVHNHSRTYNPPPYRWQDSRWGRKEL---QGTLIYQLHVGSTSTPDGTL 125
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 121 GVIRKLDYLDLGLTATEIMPIAQFPGKRDWGDVYLYAVQNSYSGSGEPGFKLVDEAHK 180
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 126 AAGEKLSYLDLGLDIFDELMPVNSFAGTRGMGVDGLVWYSHVEPYGPGDGLFRFIDAA 185
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 181 KGLGVLDVYNNHVGPGNVMWKLGPYFSQKYKTPMGLTFNFDDAESDEVKRFLEINVEY 240
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 186 AGLGVTDVYNNHVGPGNVMWKLGPYFSQKYKTPMGLTFNFDDAESDEVKRFLEINVEY 245
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 241 WIKEYNVYDGRFLDAVHAIDTSPKHILEE---TADVYHKNY-----RIVIAESLNDLPRV 294
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 246 WVGVDYHVDGVGDAVHAVERDERAVHILEDLGLADGDAISGETGLPKTLIAESDFNPRLI 305
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 295 PKEKCGYNIDAQWVDFHHISIHAYLTCERGQYTDGRLNDLDDIVKSYKDVYVQNSYSG 354
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 306 PRDVNGYGLAGOWSDDFHTAVHVSGETTYGYSDFESDLAVLAKVLKDLHGDSYSF 365
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 355 RKTGHEPVGE--LDGCNFVYIQNHQVGNRGKRIIKLVDRSYKIAAALYLLSPYIP 412
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 366 GRHGRPINSLANPAALVVCNQHNDIGNRATGDRLSQSLSYQLAVAVALTITSFPT 425
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 413 MIFMGEYGEENPFYFSDSFKLIQVREGKKE---NQD-----TDPQDESTFNASK 465
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 426 MLFMGEYGAFTPWQFTSHPELKGATAEGRIKEFERMGWDPVAVVDPDQDPTFNRSK 485
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 466 LSWK-----IDEIFSFYKILIKMRKELSTACD-----RRVNVVNGENWLLIKGREYFSL 515
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 486 LDWSEASTGDHARLLELYKSLTALRREHPLADLGLFGQTEVSDDDAGWLRFRFPVSVEVL 545
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
QY 516 YVFSKSSIEV-KYSGTLLSSNNSFP 540
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:
Db 546 VNLSDAKVRLDDAAGDLLLATDEGNP 571
||:||||| : : : : : ||:||||| : : : : : ||:||||| : : : : : ||:

RESULT 6
GLGB_BACCL STANDARD; PRT; 666 AA.
AC P30537;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE enzyme).
GN GLGB.
OS Bacillus caldolyticus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1394;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93208370; PubMed=1296817;
RX Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;
```


Query Match		12.6%; Score 374.5; DB 1; Length 627;
Best Local Similarity		23.1%; Pred. No. 3e-17;
Matches 139; Conservative		94; Mismatches 223; Indels 147; Gaps 24;
Qy	14 FTLWAPYQKSVKLVLEKGLYE-----MER-DEKGYFTITLNNVKVRDRYK-VLDD 63	
Db	1 FCVWAPASEVRV---AGDFNSGSEHVMHRVNDNGIWTLPFGICEKERYKYEIVTN 96	
Qy	64 ASEI-----PDPAS-RYQPEGVHGFSQIOESKENNETFLKEDLIIEI 108	
Db	97 NGEIRLKADPAIYAYSEVRPASTATYDLEGYSWQDKWKQKAKT---LYEKPVIYEL 153	
Qy	109 HVGFTFTPEGTEGVIRK-----LDYKLDLGTALTEIMPIAQFGPKRWDGVDGVVLAQV 162	
Db	154 HLGSKKHSDGRHSYKELSTQLTPYIKKHGFTHIELLPVYEHYDPSWGQGTGYFSPT 213	
Qy	163 NSYGGPEGFRKLVDAHKKGLGVLDVVYVNHV--GPEGNYMVKLGPFSQYKTP----- 215	
Db	214 SRFGPPHLMKFVDECHOONIGVILDWVPGHCKDAHGLYMFDEGPELYEYKEERDRENWL 273	
Qy	216 WGLTFNEDDAESDEVRKFIENVEYKYNVDGFRDVAHALI---DTSKPHILEEAD 272	
Db	274 WG-TANF-DLGGKPEVHVSFLISNALYAEFYHIDGFRDVAHALIYWPNOQDERHTNPYAVD 331	
Qy	273 VVHKYNR-----IVIAESDLNDRPVNPKCKGYNIDAQWVDPFHSHIAYLTGER 323	
Db	332 FLKLNQTMREAYPHVMIAEDSTEWQVTCGAVEGGGLGFHYKW----- 375	
Qy	324 QGYVTDFGNLDDIVKSKYDVFYDGKYSNFRKTHGEPVGLDGNFVVIQ----- 375	
Db	376 ---NMGMNDVLK-YMET-----PPEERRCHQOLISFLLYAFSEH 412	
Qy	376 ---NHQDVG-----NRCKGERIKLVDRSEYKTAALYLLSPYPMIFMGEEYGE 423	
Db	413 FVLFPSEHDEVYVYKSLNMPGSDYQKFA---QYRLLLGYMTVHPGKHLIFMGSEFAQ- 468	
Qy	424 NPFFYFSDSKTIQVREGKKGQDTPDQD-----ESTFNASKLSW--- 468	
Db	469 ---FDEKDKTEQLDWFDSFPHQKASVFTQDLLRFYQKSKLYEHDHRAQSEFIDV 523	
Qy	469 -KIDEEIFSYKILIKMRKELSIACDRVNVNGENMLIKRGYFSLVFSKSIIEVYK 527	
Db	524 HNDEQSIFSFYRQKKGAEALVIICNFTPVVYHOYD---VGVPFPTQYIEVLNSDSEY 579	
Qy	528 SGT 530	
Db	580 GGS 582	
RESULT 8		
GLGB_BACST	STANDARD;	PRT; 639 AA.
AC	P30538;	
DT	01-APR-1993 (Rel. 25, Created)	
DT	01-APR-1993 (Rel. 25, Last sequence update)	
DT	01-FEB-1995 (Rel. 31, Last annotation update)	
DE	1.4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching enzyme).	
DE	enzyme).	
GN	GLGB.	
OS	Bacillus stearothermophilus.	
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;	
OC	Bacillus/Staphylococcus group; Geobacillus.	
OX	NCBI_TaxID=1422;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=1503-4R;	
RX	MEDLINE=92079888; PubMed=1745226;	
RA	Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;	
RT	"Molecular cloning and nucleotide sequence of the glycogen branching enzyme gene (glgB) from Bacillus stearothermophilus and expression in Escherichia coli and Bacillus subtilis.";	

RL	Mol. Gen. Genet. 230:136-144(1991).
CC	-!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of glycogen.
CC	-!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC	-!- MISCELLANEOUS: THE TEMPERATURE FOR OPTIMAL ACTIVITY IS APPROXIMATELY 55 DEGREES CELSIUS.
CC	-!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC	-----
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/ or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; M35089; AAA22482.1; -
DR	PIR; S18599; S18599.
DR	InterPro; IPR00461; Alpha_amylase.
DR	InterPro; IPR004193; Isoamylase_N.
DR	Pfam; PF00128; alpha-amylase_N.
DR	Pfam; PF02922; isoamylase_N; 1.
KW	Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT	ACT_SITE 309 309 BY SIMILARITY.
FT	ACT_SITE 352 352 BY SIMILARITY.
FT	ACT_SITE 420 420 BY SIMILARITY.
SQ	SEQUENCE 639 AA; 74795 MW; F5BD4446B371E03A CRC64;

Qy	4 AYKIDGNEVI-----FTLWAPYQKSVKL-----KVLKGLYEMERD 39	
Db	22 SYELFGAHVIKKNMGVGRFCVWAPHAREVLGVSFNENGTNFMKYSQGVMMIP 81	
Qy	40 E-KGYFTITLNNVKVRDRYKVL--DDASEI--PDASRYQEGVHGSPSIIQ-ESKEF 92	
Db	82 ENLSEGL-----YKEITNDGCVLLKSDPYAFYSELPHPTASIVYIKGQW 129	
Qy	93 NNETFLKK-----EDLIIEIHVGTF--TPGTF---EGVIRKLDYIKDGLITAIE 138	
Db	130 NDQWRRKKQRKRIYDQPLFIYELHFGSKKKEDSGSFYQEMAEELIPYVLEHGFTHIE 189	
Qy	139 IMPAQPGKRDWGDGYLYAVQNSYCGPEGFRKLVDEAHKGLGVLDVVYVNHV--GP 196	
Db	190 LLPVHEHPFDRSWGQIGYYSATSRVGTPHDLMFYIDRCHOAGIGVILDWVPGHCKDS 249	
Qy	197 EGNVMVKLGYPF-----SQYKTPWGLTFNFDAAESDEVKRFILNEYWIKEYNVDCFR 251	
Db	250 HGLYFDCAPAYEYANQDRENYVWG-TANF-DLGKPEVRSFLISNALFWMEYFHVDFGR 307	
Qy	252 LDVAHAI-----DTSKPHILEIADVHKY--NRVIAESDLNDRPVNPKKCK 299	
Db	308 VDAVANMLYWPNSDVLVKNYAVEFLQKLNETVAYDNPILMAIEDSTDWPRVTAPTYDG 367	
Qy	300 GYNDAQWVDFFHSIHAYLTGERQGYTDFGNLDDIVKSKYDVFYDGKYSNFRKTHG 359	
Db	368 GLGFNYKW-----NMGMNDIL-----TYMETPEHRKYVHN 399	
Qy	360 EPVGELDGCNFWVIYQN-----HDQVGNRGKGERIKLVDRSEYKTAAL-YYYLS- 408	
Db	400 KVTESL-----LYAYSENFLPFSHDEVVH-GKSKLSKMPCTYEKPAQLRLLYGLLTH 454	
Qy	409 PYIPMIFMGEEYGENPYFSDSFKLIQGVREGKKGQDTPDDESTFNASKLSW 468	
Db	455 PGKLLFMGGEFGQ-----FDEWKD-----LEQLDW 480	
Qy	469 KIDEEIFSYKILIKMRKELSIACDRRVNVV-----NGENWLLIKRGYFSLYVF 518	
Db	481 MLFD--FDMHRNMNMYVKEL-LKCYKRYKPLYELDHDSPDFEWDVHNAEQ-SIFS 533	

Query Match	12.6%; Score 374.5; DB 1; Length 639;
Best Local Similarity	23.6%; Pred. No. 3e-17;
Matches 141; Conservative	99; Mismatches 190; Indels 167; Gaps 29;


```
CC EMBL; L11647; AAA67437.1; -
DR InterPro; IPR000461; Alpha_amyase.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amyase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT ACT_SITE 440 440 BY SIMILARITY.
FT ACT_SITE 493 493 BY SIMILARITY.
FT ACT_SITE 561 561 BY SIMILARITY.
SQ SEQUENCE 764 AA; 85325 MW; 6B45482EAA268ACF CRC64;

Query Match 11.2%; Score 335; DB 1; Length 764;
Best Local Similarity 24.2%; Pred. No. 1.5e-14;
Matches 133; Conservative 90; Mismatches 137; Indels 130; Gaps 26;

Qy 14 FTLWAPYQKSVKL---KVLEKGLYMER-DEKGYFTITLNNVVRDRIKYL---DDAS 65
Db 178 FTVWAPNALGVYRVGTGDFSYWDVAVPMKSLGASGWMELFPGVAGALYKYEITRPDGR 237
Qy 66 EI-PDPASRYQEGVHGVSQIQESKEFNNETFLKKEDLI-----IYEHVGTFTPE 116
Db 238 TLRADPMARYAEVPPANASIVTASRYEQDAEWARRGALAPHQPMVSVEELHLSWRPG 297
Qy 117 GTFEGVIRKLD-YLKDLGITAIEIMPVIAOFFGKRDMGVDGYLVAVQNSYGGPEGFRKL 175
Db 298 LSYRQLAEQLPAYVKELGFTHVLELMPVAEHPFGGSGVQVTFGYAPTSRMGTDPDFRLV 357
Qy 176 DEAKKGLGVLDVYVNVHVGPEGNYVWKLGPYFSQKYTP-----WGLTFNDDAESD 228
Db 358 DALHRAGIGVIVDVPVPAHFPRDDWALAEFDGRPLYEHQDPRRAHPDMWG-TLEFDYGRK 415
Qy 229 EVRKFILENVYWKYKYNVDGFRLDVAHAI-----DTSPK-----HTLE 268
Db 416 EVRNLVANAYWCQEFHVDGLRADAVASMLYDYSRDEGDSWNAHGREDLDAVALQ 475
Qy 269 EI-ADVHKYNRIV-IAESLDLPRVNPKEKCGYNIDQAVDDFHHSIAHYLGE---R 323
Db 476 EMNATVYRRFPVGVVTIABESTAWDGVTRPTDSGGGLGKWNMGWMDTLRYVSKPEVHR 535
Qy 324 QGYTID-FGNLDDIVKSYKDVYVDGKYSNFRKTHGEPVGLDGCNFFVYVIONHDQVG 381
Db 536 KYHHDMTFG---WVAFSE-----NFVLPV-SHDEV 564
Qy 382 NRGKGERIKL-----VDRESYKIAAALYLLSPYIPMFMEGEY--GEE-----NPYFF 429
Db 565 H-GRSLVSKMPGDWQWQORATHRAYLGFPMWHPCKQLLFMGQEFQAQGSSESETYGPDMW 623
Qy 430 SDFS-----DSKLIQGVREGKKENGQDTPQDESTFNASKLSWK----- 469
Db 624 LDSSYPAAGDH---LGVRSRLVDLN-----RTVTASPALWERDSVPBGFAWEADA 671
Qy 470 IDEEIFSYK 479
Db 672 ADDNVFAFLR 681

RESULT 11
GLGB_SYNY3
ID GLGB_SYNY3 STANDARD; PRT; 770 AA.
AC P32981;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE enzyme).
GN GLGB OR SLL0158.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugitara M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -|- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC glycogen.
CC -|- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC -|- SUBUNIT: MONOMER (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D63999; BAA10073.1; -
DR InterPro; IPR000461; Alpha_amyase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amyase; 1.
DR Pfam; PF02922; isoamylase_N; 2.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW Complete proteome.
FT ACT_SITE 433 433 BY SIMILARITY.
FT ACT_SITE 486 486 BY SIMILARITY.
FT ACT_SITE 554 554 BY SIMILARITY.
SQ SEQUENCE 770 AA; 89527 MW; A435AFCA7703FA8A CRC64;

Query Match 11.2%; Score 333.5; DB 1; Length 770;
Best Local Similarity 23.1%; Pred. No. 1.9e-14;
Matches 119; Conservative 91; Mismatches 176; Indels 129; Gaps 20;

Qy 7 IDG-NEVIFTLWAPYQKSVKL---KVLEKGLYMERDEKGYFTITLNNVVRDRIKYL 61
Db 136 VDGKGYFFAWAPNARNVSIILGDFNNDGLHQMRRNNMWELFPELGVGYSYKYEI 195
Qy 62 DD-----ASEIPDPASRYQEGVHGVSQIQ-ESKEFNNETFLK-----KEDLIYEI 108
Db 196 KNEGHIVTEKTDYGFYQEVPRKTAIVADLDGQWHDWLEARRTSDDLSPKSPVVEL 255
Qy 109 HVG-----TFTPECT-----FEGVIRKLDYLDLGLTATEIM 140
Db 256 HLGSLWHTAYDEPVKTLHGECPVVEVSEWNTGARFLTYVELVDKLIIPYKELGYTHIEL 315
Qy 141 PIAQFPKRDWGYGVLYAVONSYGGPGEGRKLVDAHKKGLGVLDVYVNVHVGPEGN- 199
Db 316 PIAEHPDGSWGYQVYGYAPTSPFGSPEDPMYFVDOCHLNGIIVDWPDPKDHG 375
Qy 200 -----YMKLGPYFSQKYTPWG-LTFNPDAADESDEVRKFEILENVEYWKYKYNVDGFR 251
Db 376 LAFFDGTGLYEHGDPKRGHEK-ENGTLIFNYG---RNEVRNLFVANALFWFDKVIDGMR 431
Qy 252 LDVAHAI-----DTSPKHILIEIADVHKY--NRIVIAESLDN 289
Db 432 VDAVASMLYLDYCREGEWANEYGGRENLEAADFLLQVNSVWYSYFPGILSIAESTSW 491
Qy 290 PRVNPKEKCGYNIDQAVDDFHHSIAHYLTGERQYVTFDGNLDDIVKSYKDVGVVDGK 349
Db 492 PWSWPTVYVGLGNLKW-----NWGMHMDML-----DYFSMDPW 526
Qy 350 YSNFRKT-----HGEVPGVGLDGCNFFVYVIONHDQVGNRGKGERIKLVDRSEYKIA 401
Db 527 FRQPHQNSITFSWMYHNSH-----NYMLAL-SHDEVVH-GKSNMLGKMPGDEWKYA 576
Qy 402 AALYLLS-----PYIPMFNGEEYGEENPFYFFSD 431
```



```

Query Match      10.7%  Score 318;  DB 1;  Length 659;
Best Local Similarity 23.1%;  Pred. No. 1.6e-13;
Matches 137;  Conservative 98;  Mismatches 173;  Indels 186;  Gaps 34;

QY  3  FAYKIDG-----NEVIFTLWAPQKSKYKLVKLEKGLYMERDE-KGYFTITLNNKVRDR 56
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  74  YAFRIHGEFANPKLIID---PYAKAY-----NGKPOLSESKRSKFLLISDN-----RDN 120

```

Search completed: July 15, 2002, 12:24:32
Job time: 384 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:17:38 ; Search time 55.94 Seconds
(without alignments)
1728.712 Million cell updates/sec

Title: US-09-298-924-6
Perfect score: 2978
Sequence: 1 MTFAYKIDGNEVFTLWAPY.....PQHIEGKYEFDFKGALYKL 559

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2978	100.0	559	1 Q55088	Q55088 sulfolobus
2	2837	95.3	559	1 Q9UWN9	Q9UWN9 sulfolobus
3	2467	82.8	561	17 P95867	P95867 sulfolobus
4	1748.5	58.7	556	1 Q53641	Q53641 sulfolobus
5	1269	42.6	359	17 Q973H4	Q973H4 sulfolobus
6	1033.5	34.7	581	2 Q9ADI5	Q9ADI5 streptomyces
7	905.5	30.4	600	16 Q9RX51	Q9RX51 deinococcus
8	881.5	29.6	583	16 Q9I1V1	Q9I1V1 pseudomonas
9	876.5	29.4	601	16 Q92U63	Q92U63 rhizobium m
10	706	23.7	217	17 Q973H5	Q973H5 sulfolobus
11	524.5	17.6	552	2 Q44528	Q44528 anabaena va
12	399	13.4	652	2 Q59242	Q59242 bacillus st
13	397	13.3	440	2 Q93Q35	Q93Q35 myxococcus
14	395.5	13.3	1142	2 Q930Z9	Q930Z9 bacillus sp
15	381.5	12.8	825	2 Q59319	Q59319 caldocellum
16	365	12.3	1280	16 Q97SQ7	Q97SQ7 streptococc

17	364	12.2	1287	2	Q9F930	Q9F930 streptococc
18	361	12.1	630	16	O66936	O66936 aquifex aeo
19	349.5	11.7	783	10	O04196	O04196 arabidopsis
20	348.5	11.7	666	16	O84046	O84046 chlamydia t
21	346.5	11.6	713	1	O05152	O05152 sulfolobus
22	345.5	11.6	422	2	O59243	O59243 bacillus st
23	345.5	11.6	664	16	Q928F5	Q928F5 chlamydia p
24	345.5	11.6	666	16	Q9PKZ6	Q9PKZ6 chlamydia m
25	345.5	11.6	707	16	P73608	P73608 synechocyst
26	344.5	11.6	741	2	Q9KV06	Q9KV06 streptomyces
27	344.5	11.6	741	2	Q59832	Q59832 streptomyces
28	343	11.5	718	17	P95868	P95868 sulfolobus
29	341	11.5	737	16	Q985P4	Q985P4 rhizobium 1
30	339.5	11.4	789	10	O22637	O22637 zea mays (m
31	336	11.3	702	10	Q9M0S5	Q9M0S5 arabidopsis
32	336	11.3	716	17	Q973H3	Q973H3 sulfolobus
33	335	11.2	1072	16	Q9K6N1	Q9K6N1 bacillus ha
34	334.5	11.2	717	16	Q9K7U5	Q9K7U5 bacillus ha
35	334.5	11.2	736	16	Q92M14	Q92M14 rhizobium m
36	333.5	11.2	720	16	Q9RXP5	Q9RXP5 deinococcus
37	331.5	11.1	1938	2	P70983	P70983 bacillus sp
38	330.5	11.1	562	2	Q9L872	Q9L872 bacillus sp
39	330.5	11.1	562	16	Q9K7J1	Q9K7J1 vibrio chol
40	329.5	11.1	668	2	P71095	P71095 bacteroides
41	329.5	11.1	818	10	Q41742	Q41742 zea mays (m
42	329.5	11.1	1165	16	Q99XX8	Q99XX8 streptococc
43	329	11.0	741	3	Q9P5P3	Q9P5P3 neurospora
44	327	11.0	733	10	O80403	O80403 oryza sativ
45	323.5	10.9	718	16	O34587	O34587 bacillus su

ALIGNMENTS

RESULT	1
Q55088	PRELIMINARY; PRT; 559 AA.
AC	Q55088;
DT	01-NOV-1996 (TREMREL. 01, Created)
DT	01-MAY-1997 (TREMREL. 03, Last sequence update)
DT	01-DEC-2001 (TREMREL. 19, Last annotation update)
DE	ALPHA-AMYLASE.
OS	Sulfolobus solfataricus.
OC	Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX	NCBI_TaxID=2287;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=KMI;
RA	Kato M., Kettoku M., Miura Y., Komeda T., Konishi Y., Shindo K.,
RA	Kobayashi K., Iwanatsu A.;
RT	"The gene analysis of the new amylases from the hyper thermophilic
RT	archae Sulfolobus";
RL	Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
DR	EMBL; D64130; BAA1010.1; -.
DR	InterPro; IPR00461; Alpha-amylase.
DR	InterPro; IPR004193; isoamylase N.
DR	Pfam; PF00128; alpha-amylase; 1.
DR	Pfam; PF02922; isoamylase-N; 1.
SQ	SEQUENCE 559 AA; 64790 MW; 79FBE23A7CD38B4E CRC64;

Query Match	100.0%;	Score 2978;	DB 1;	Length 559;
Best Local Similarity	100.0%;	Pred. No. 1.3e-180;		
Matches 559;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MTFAYKIDGNEVFTLWAPYQKSVKLKLEKGLYEMERDEKGYFTITLNNKVRDRYKYV	60	
Db	1	MTFAYKIDGNEVFTLWAPYQKSVKLKLEKGLYEMERDEKGYFTITLNNKVRDRYKYV	60	
QY	61	LDASIPDPASRYQEGVHGPSQIQESKEFNNEFLKEDLLIYEHVGTFTPGTPE	120	
Db	61	LDASIPDPASRYQEGVHGPSQIQESKEFNNEFLKEDLLIYEHVGTFTPGTPE	120	

```
OY 121 GVIRKLDYLKDLGITAIEIMPIAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
Db 121 GVIRKLDYLKDLGITAIEIMPIAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
OY 181 KGLGVILDVYVNHVGPENYMWKLGYPFSQKYKTPWGLTFNFDAAESDEVKRFILLENVEY 240
Db 181 KGLGVILDVYVNHVGPENYMWKLGYPFSQKYKTPWGLTFNFDAAESDEVKRFILLENVEY 240
OY 241 WKEYNVNVDGFRDVAHAIIDTSPKHILEEADVHVHKNRIVIAESDLNDPRVVPKKEG 300
Db 241 WKEYNVNVDGFRDVAHAIIDTSPKHILEEADVHVHKNRIVIAESDLNDPRVVPKKEG 300
OY 301 YNIDAQWDDFHHSIHAYLTGERQGYDYGNDLDDIVKSYKDVYDGYKSNFRKKTG 360
Db 301 YNIDAQWDDFHHSIHAYLTGERQGYDYGNDLDDIVKSYKDVYDGYKSNFRKKTG 360
OY 361 PVGELDGCNFVYVYQNHQDVGNRGKGERIKLVDRSYKIAAALYLLSPYIPMIPMGREY 420
Db 361 PVGELDGCNFVYVYQNHQDVGNRGKGERIKLVDRSYKIAAALYLLSPYIPMIPMGREY 420
OY 421 GEENPFYFFSDSKLIQGVREGKKGQDTPQDESTFNASKLSWKIDEEIFSFKI 480
Db 421 GEENPFYFFSDSKLIQGVREGKKGQDTPQDESTFNASKLSWKIDEEIFSFKI 480
OY 481 LIKMKKELSIACDRRVNVVNGENWLLIKGREYFSLYVFSKSIEVKYSGTLLSSNNF 540
Db 481 LIKMKKELSIACDRRVNVVNGENWLLIKGREYFSLYVFSKSIEVKYSGTLLSSNNF 540
OY 541 QHIEGKYEFKGFALYKL 559
Db 541 QHIEGKYEFKGFALYKL 559

RESULT 2
OY 121 GVIRKLDYLKDLGITAIEIMPIAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
Db 121 GVIRKLDYLKDLGITAIEIMPIAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
OY 181 KGLGVILDVYVNHVGPENYMWKLGYPFSQKYKTPWGLTFNFDAAESDEVKRFILLENVEY 240
Db 181 KGLGVILDVYVNHVGPENYMWKLGYPFSQKYKTPWGLTFNFDAAESDEVKRFILLENVEY 240
OY 241 WKEYNVNVDGFRDVAHAIIDTSPKHILEEADVHVHKNRIVIAESDLNDPRVVPKKEG 300
Db 241 WKEYNVNVDGFRDVAHAIIDTSPKHILEEADVHVHKNRIVIAESDLNDPRVVPKKEG 300
OY 301 YNIDAQWDDFHHSIHAYLTGERQGYDYGNDLDDIVKSYKDVYDGYKSNFRKKTG 360
Db 301 YNIDAQWDDFHHSIHAYLTGERQGYDYGNDLDDIVKSYKDVYDGYKSNFRKKTG 360
OY 361 PVGELDGCNFVYVYQNHQDVGNRGKGERIKLVDRSYKIAAALYLLSPYIPMIPMGREY 420
Db 361 PVGELDGCNFVYVYQNHQDVGNRGKGERIKLVDRSYKIAAALYLLSPYIPMIPMGREY 420
OY 421 GEENPFYFFSDSKLIQGVREGKKGQDTPQDESTFNASKLSWKIDEEIFSFKI 480
Db 421 GEENPFYFFSDSKLIQGVREGKKGQDTPQDESTFNASKLSWKIDEEIFSFKI 480
OY 481 LIKMKKELSIACDRRVNVVNGENWLLIKGREYFSLYVFSKSIEVKYSGTLLSSNNF 540
Db 481 LIKMKKELSIACDRRVNVVNGENWLLIKGREYFSLYVFSKSIEVKYSGTLLSSNNF 540
OY 541 QHIEGKYEFKGFALYKL 559
Db 541 QHIEGKYEFKGFALYKL 559

RESULT 2
ID Q9UNW9 PRELIMINARY; PRT; 559 AA.
AC Q9UNW9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MALTOOLIGOSYL TREHALOSE TREHALOHYDROLASE.
GN TREZ.
OS Sulfolobus shibatae.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2286;
RN [1]
RP SEQUENCE FROM N.A.
RA Weiting Y., Wei C., Hui W., Li L., Cheng J.;
RT "Maltooligosyl trehalose trehalohydrolase from Sulfolobus shibatae.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF201335; AAF17553.1; -
DR InterPro; IPR000461; Alpha_amyase.
DR InterPro; IPR004193; Isoamylase_N.
DR Pfam; PF00128; alpha-amyase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Hydrolase.
SQ SEQUENCE 559 AA; 64619 MW; 8C0F0F38A5AE2846 CRC64;

Query Match 95.3%; Score 2837; DB 1; Length 559;
Best Local Similarity 95.2%; Pred. No. 1.1e-171;
Matches 532; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

OY 1 MTFAYKIDGNEVIFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYV 60
Db 1 MTFAYKIDGNEVIFTLWAPYQKSVKLVLEKGLYEMERDDKGYFTITLNNVKVDRYKYV 60

OY 61 LDDASEIPDPASRYQPGVGPSPQIIQESKEFNNETLKKEDLIYIHHVGTFTPEGTTF 120
Db 61 LDDASEIPDPASRYQPGVGPSPQIIQESKEFNNETLKKEDLIYIHHVGTFTPEGTTF 120

OY 121 GVIRKLDYLKDLGITAIEIMPIAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
Db 121 GVIRKLDYLKDLGITAIEIMPIAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
```

```
Db 121 GVIRKLDYLKDLGITAIEIMPIAQPPGKRDWGYDGYLYAVONSYGGPGEFRKLVDEAHK 180
OY 181 KGLGVILDVYVNHVGPENYMWKLGYPFSQKYKTPWGLTFNFDAAESDEVKRFILLENVEY 240
Db 181 KGLGVILDVYVNHVGPENYMWKLGYPFSQKYKTPWGLTFNFDAAESDEVKRFILLENVEY 240
OY 241 WKEYNVNVDGFRDVAHAIIDTSPKHILEEADVHVHKNRIVIAESDLNDPRVVPKKEG 300
Db 241 WKEYNVNVDGFRDVAHAIIDTSPKHILEEADVHVHKNRIVIAESDLNDPRVVPKKEG 300
OY 301 YNIDAQWDDFHHSIHAYLTGERQGYDYGNDLDDIVKSYKDVYDGYKSNFRKKTG 360
Db 301 YNIDAQWDDFHHSIHAYLTGERQGYDYGNDLDDIVKSYKDVYDGYKSNFRKKTG 360
OY 361 PVGELDGCNFVYVYQNHQDVGNRGKGERIKLVDRSYKIAAALYLLSPYIPMIPMGREY 420
Db 361 PVGELDGCNFVYVYQNHQDVGNRGKGERIKLVDRSYKIAAALYLLSPYIPMIPMGREY 420
OY 421 GEENPFYFFSDSKLIQGVREGKKGQDTPQDESTFNASKLSWKIDEEIFSFKI 480
Db 421 GEENPFYFFSDSKLIQGVREGKKGQDTPQDESTFNASKLSWKIDEEIFSFKI 480
OY 481 LIKMKKELSIACDRRVNVVNGENWLLIKGREYFSLYVFSKSIEVKYSGTLLSSNNF 540
Db 481 LIKMKKELSIACDRRVNVVNGENWLLIKGREYFSLYVFSKSIEVKYSGTLLSSNNF 540
OY 541 QHIEGKYEFKGFALYKL 559
Db 541 QHIEGKYEFKGFALYKL 559

RESULT 3
ID P95867 PRELIMINARY; PRT; 561 AA.
AC P95867;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE PRECURSOR (MALTO-OLIGOSYLTREHALOSE TREHALOHYDROLASE)
DE (TREZ) (EC 3.2.1.141).
GN TREZ.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Wayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; Y08256; CAA69503.1; -
DR EMBL; AE006815; AAK42272.1; -
DR InterPro; IPR000461; Alpha_amyase.
DR InterPro; IPR004193; Isoamylase_N.
DR Pfam; PF00128; alpha-amyase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Hydrolase; Glycosidase; Complete proteome.
SQ SEQUENCE 561 AA; 64370 MW; B00EA03020F6B242 CRC64;

Query Match 82.8%; Score 2467; DB 17; Length 561;
Best Local Similarity 79.7%; Pred. No. 3e-148;
Matches 447; Conservative 63; Mismatches 49; Indels 2; Gaps 1;

OY 1 MTFAYKIDGNEVIFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYV 60
Db 1 MTFAYKIDGNEVIFTLWAPYQKSVKLVLEKGLYEMERDEKGYFTITLNNVKVDRYKYV 60
```

Db 1 MTFGKLEDDGVTFNLWAPYQKVKLILNRGIYEMERDDKGYFTITLDNVRVGDYKYI 60
QY 61 LDADSEIPDPASRYQPGVHGSPSOIQESKEFNNEF--LKKEDELIYETHVGTFTPEGT 118
Db 61 LDADSEIPDPASRYQPGVHGSEITSPDEFWDDNSVKYKREDLVYIELHIGFTFSEGT 120
QY 119 FEGVIRKLDYKDLGITAIEIMPIAOPGKRDGMDGYVLYAVQNSYGGPGRKFLVDEA 178
Db 121 FEGVIRKLDYKDLGITAIEIMPIAOPGKRDGMDGYVLYAVQNSYGGPGRKFLVDEA 180
QY 179 HKKGLGLVLDVYVNHVGPENYVVKLGYPYSQKYKTPWGLTFNFDDEADSEVRFKFIENV 238
Db 181 HKKGLGLVLDVYVNHVGPENYVVKLGYPYSQKYKTPWGLTFNFDDEADSEVRFKFIENV 240
QY 239 EYWKIKEYNDGFRDLDAVHAIDTSPKHILEEIAADVHKYNRIVIAESDLNDPRVNPKEK 298
Db 241 EYWKIKEYNDGFRDLDAVHAIDTSPKHILEEIAADVHKYNRIVIAESDLNDPRVNPKEK 300
QY 299 CGYNIDAQWDDFHHSTHAYLTGEROGYTDGFLNDIDVKSVDYVYDGYKYSNFRKTH 358
Db 301 CGYNIDAQWDDFHHSTHAYLTGEROGYTDGFLNDIDVKSVDYVYDGYKYSNFRKTH 360
QY 359 GEPVGLDGCNFVYVYQNHQDQVGNRGKGERIILKLVDRSEYKIAAALYLLSPYIPMIFMGE 418
Db 361 GKSVDLDGCKFVYVYQNHQDQVGNRGKGERIILKLVDRSEYKIAAALYLLSPYIPMIFMGE 420
QY 419 EYGEENPFYFSDFSKLIQGVREGKKGQDTPDQESTFNASKLSWKIDEEIFSFY 478
Db 421 EYGEENPFYFSDFSKLIQGVREGKKGQDTPDQESTFNASKLSWKIDEEIFSFY 480
QY 479 KILIKMRKELSIACDRRVNVNGENMLIIKRGREYFSLYVFSKSYEGYSGTLLLSNNS 538
Db 481 KSLIKIRKEYGLACNRKLSVNGENYVLTGKNGCLAVYVFSKSYEGYSGTLLLSNNS 540
QY 539 FPOHIEGKGFYFSDKGFALYKL 559
Db 541 FPOHIEGKGFYFSDKGFALYKL 561

RESULT 4
Q53641
ID Q53641 PRELIMINARY; PRT; 556 AA.
AC Q53641; 008279; 008064;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE (FRAGMENT).
GN TREZ.
OS Sulfolobus acidocaldarius.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2285;
RN [1]
RP SEQUENCE FROM N.A.
RA Kato M., Kettoku M., Miura Y., Komeda T., Konishi Y., Shindo K.,
RA Kobayashi K., Iwanatsu A.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC33909;
RX MEDLINE=97135071; PubMed=8980629;
RA Maruta K., Mitsuzumi H., Nakada T., Kubota M., Chaen H., Fukuda S.,
RA Sugimoto T., Kurimoto M.;
RT "Cloning and sequencing of a cluster of genes encoding novel enzymes
RT of trehalose biosynthesis from thermophilic archaeobacterium Sulfolobus
RT acidocaldarius.";
RL Biochim. Biophys. Acta 1291:177-181(1996).
DR EMBL; D64131; BAA11011.1; -;
DR EMBL; D83245; BAA11863.1; -;
DR InterPro; IPR000461; Alpha.amylase.
DR InterPro; IPR0004193; isoamylase.N.
DR InterPro; IPR003682; sub.transporter.
DR Pfam; PF00128; alpha.amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
KW Hydrolase.
FT NON_TER 556
SQ SEQUENCE 556 AA; 64373 MW; 4215B45E6C8E04E7 CRC64;

Query Match 58.7%; Score 1748.5; DB 1; Length 556;
Best Local Similarity 58.9%; Pred. No. 8.9e-103;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

QY 2 TFAYKIDGNEVITLWAPYQKVKLILNRGIYEMERDDKGYFTITLDNVRVGDYKYI 61
Db 3 SFGNIEKNKGIKFLWAPYVNSVKL--LSKKLIPMEKNDEGFEDEITLLEENLYSII 61
QY 62 DDAASEIPDPASRYQPGVHGSPSOIQESKEFNNEF--LKKEDELIYETHVGTFTPEGT 121
Db 62 DDAASEIPDPASRYQPGVHGSPSOIQESKEFNNEF--LKKEDELIYETHVGTFTPEGT 121
QY 122 VIRKLDYKDLGITAIEIMPIAOPGKRDGMDGYVLYAVQNSYGGPGRKFLVDEAHHK 181
Db 122 VIRKLDYKDLGITAIEIMPIAOPGKRDGMDGYVLYAVQNSYGGPGRKFLVDEAHHK 181
QY 182 GLGVLDVYVNHVGPENYVVKLGYPYSQKYKTPWGLTFNFDDEADSEVRFKFIENV 241
Db 182 GLGVLDVYVNHVGPENYVVKLGYPYSQKYKTPWGLTFNFDDEADSEVRFKFIENV 241
QY 242 IKENYNDGFRDLDAVHAIDTSPKHILEEIAADVHKYNRIVIAESDLNDPRVNPKEKCGY 301
Db 242 IKENYNDGFRDLDAVHAIDTSPKHILEEIAADVHKYNRIVIAESDLNDPRVNPKEKCGY 301
QY 302 NIDAQWDDFHHSTHAYLTGEROGYTDGFLNDIDVKSVDYVYDGYKYSNFRKTHGEP 361
Db 302 NIDAQWDDFHHSTHAYLTGEROGYTDGFLNDIDVKSVDYVYDGYKYSNFRKTHGEP 361
QY 362 VGELDGCNFVYVYQNHQDQVGNRGKGERIILKLVDRSEYKIAAALYLLSPYIPMIFMGE 421
Db 362 VGELDGCNFVYVYQNHQDQVGNRGKGERIILKLVDRSEYKIAAALYLLSPYIPMIFMGE 421
QY 422 EENPFYFSDFSKLIQGVREGKKGQDTPDQESTFNASKLSWKIDEEIFSFYKIL 481
Db 422 EENPFYFSDFSKLIQGVREGKKGQDTPDQESTFNASKLSWKIDEEIFSFYKIL 481
QY 482 IKMRKELSIACDRRVNVNGENMLIIKRGREYFSLYVFSKSYEGYSGTLLLSNNSFPQ 541
Db 482 IKMRKELSIACDRRVNVNGENMLIIKRGREYFSLYVFSKSYEGYSGTLLLSNNSFPQ 541
QY 542 HIEGK-YEFDKGFALYKL 559
Db 542 HIEGK-YEFDKGFALYKL 559
QY 559 HIEGK-YEFDKGFALYKL 559
Db 559 HIEGK-YEFDKGFALYKL 559

RESULT 5
Q973H4
ID Q973H4 PRELIMINARY; PRT; 359 AA.
AC Q973H4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE ALPHA-AMYLASE.
GN ST0927.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamaya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kushiya N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;

DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase N; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 600 AA; 66909 MW; 594091EC093F8A44 CRC64;

Query Match 30.4%; Score 905.5; DB 16; Length 600;
Best Local Similarity 38.7%; Pred. No. 2.3e-49;
Matches .205; Conservative 86; Mismatches 170; Indels 69; Gaps 20;

QY 14 FTLWAPYQKSVKLVK-----LEKGLYMERDEKGYFTITLNNVYVDRYKYLDDA 64
DB 38 FLWTSTARTAVRVNGTEHVMTSLGGIYELEP-----VGPGARYLFLVDGV 86
QY 65 SELPDPASRYQPGVHGSPQIIQESKEFNNEFR-----LKEDLIIEIYHVGTFP 115
DB 87 -PTPDYARELPDGVHGEAEV---DFG--TFDWTADWHGKIKLADCFYEVHVGTFP 139
QY 116 EGTFFGVIRKLDYKLDGITAIEIMPIAOPFGKRDNGYDGVLYAVONSYGGPEGFRKL 175
DB 140 EGYTAAAEKLPYLKELGVTAIQVMPLAAPDQGRGWYDGAFAFYAPYGRPEDLMALV 199
QY 176 DEAHKKGVLGVLDVYVNHVCPGNYVWKLGP-YFSOKYKTPWGLTFNFDDAESDEVKFI 234
DB 200 DAAHRLGLGVLDVYVNHVCPGNYLSSYAPSYFTDRFSSAWGM--GLDYAE-PHMRRY 256
QY 235 LENVEYWIKEYNVDGRLDAVHAIDTSPKHILEEIAADVVKY--NRIVIAESDLNDPRV 292
DB 257 TGNARWLRDYHEDGLRLDTPYTDSDSETHTLTAQELHGLGTHLLAEDHRLPDL 316
QY 293 VNPKEKGYNIDAOVDDPHSHYLTGERQGYTDF-GNLDIDVYKVDVYVQKYS 351
DB 317 VTYN-----HLDGITWDDFHHETRVLTGTQEGYAGYRGGAALAYTTRRGWRYEGQFW 371
QY 352 NFRKTH--GEPVGLDGCNFVYVYIONHDQVGNRGKGERIKL--VDRESYKIAAALYLL 407
DB 372 AVAGEEHEHGPSDALEAPNFYVYICQNDQIGNRPLGERLHOSDGYTLHEYRGAAALL- 430
QY 408 SPYIPMIFMGEYGEENPFYFFSDSKLIQVREGKRKE-----NGQDT-DPQDES 459
DB 431 -PMTPLLFQGEWAASTPQFFSDHA-GELGQAVSEGRKKEFGFGSGEDVPDPAEQ 488
QY 460 TFNASKLSKIDE-----EIFSYKILIKMKELSTACDR-RVNVYNGEN 503
DB 489 TFLNSKLNLAEREGGHARTLRYDRLRLRRDPVLRNQRLENLTGHD 538

RESULT 8

Q911V1 PRELIMINARY; PRT; 583 AA.
AC Q911V1;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE PROBABLE GLYCOSYL HYDROLASE.
GN PA2164.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saiter M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).

DR EMBL; AE004643; AAG05552.1; --
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 583 AA; 65657 MW; 651CD7A918833E00 CRC64;

Query Match 29.6%; Score 881.5; DB 16; Length 583;
Best Local Similarity 36.3%; Pred. No. 7.3e-48;
Matches .203; Conservative 95; Mismatches 174; Indels 87; Gaps 19;

QY 3 FAYKIDGN-EVIFTLWAPYQKSVKLVKELGYMERDEKGYFTITLNNVYVDRYKYL 61
DB 7 FGAQFOGNGRTCEGLWAPDAREVRVETADGRWPLSGSDSGWFEATL-PCPCTRYRYRI 65
QY 62 DDASETPDASRYQPGVHGSPQIIQES-----KEFNETFLKKEDLIIEIYHVGTFP 115
DB 66 DGRPGVPDPASQFQPDGVHGSQVLDHGTYAMRVDEWGRPW---HEAVIYELHVLGF-- 120
QY 116 EGTFFGVIRKLDYKLDGITAIEIMPIAOPFGKRDNGYDGVLYAVONSYGGPEGFRKL 175
DB 121 -GSYAEVERPLRLVELGVTAVELMPLGEPGRRNGYDGVLPFAPASAYGTPEQLKHLI 179
QY 176 DEAHKKGVLGVLDVYVNHVCPGNYVWKLGP-YFSOKYKTPWGLTFNFDDAESDEVKFI 234
DB 180 DSAHGGLAMVFDVYVNHVCPGNYLAQYAAAFRRDROTQWQAIDF---RRGEVREFF 236
QY 235 LENVEYWIKEYNVDGRLDAVHAIDTSPKHILEEIA-----DVVHYKNRIVIAE 284
DB 237 YENALMWLLDYRVDGLRFDVAHAIPDSA---FLVEMARRLRGAAGPERHVH-----LVLE 288
QY 285 SLDLNDPRVNPKEKGYNIDAOVDDPHSHYLTGERQGYTDFGN-LDDIVKSYKDV 343
DB 289 NDNRRASLL---RQGY--DAQNDDGHALVLLTGENDGYFYQDYPEPLRCLARCLAG 342
QY 344 FVYDGYSNFRRTKTHGEPVGLDGCNFVYVYIONHDQVGNRGKGERIKLVDRESYKIAA 403
DB 343 FVYQGE-ANRHGRPRCEPSADLAPDAFVLFQNHQDQVGNRATGERLSVLAEPQALRLA 401
QY 404 LYLLSYIPMIFMGEYGEENPFYFFSDSKLIQVREGKRKEGQD----- 452
DB 402 LQALLPMPILFLFMGECAAREPFLYETD-HQGLADAVREGRKKEFGFEGFEGATLAS 460
QY 453 -TDPQDESTFNASK-----LSWKIDEEIFSYKILIKMKELST----- 490
DB 461 LPDPNAVETFRSRPGLADCDPAWR-----GFYROLLEIRHEHLIPRLRGARSLGVTTI 514
QY 491 ---ACDRRVNVVNGENWLI 506
DB 515 AGAALSARWRLGDSWDRI 533

RESULT 9

Q92U63 PRELIMINARY; PRT; 601 AA.
AC Q92U63;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PUTATIVE 1,4-ALPHA-GLUCAN BRANCHING ENZYME PROTEIN (EC 2.4.1.18).
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,


```
Db 64 EEDQVVDTPYATDIDES-SGKDNSIARIKDGKEIKVDTYVMQHDHDKPLPADHELVIYEL 122
QY 109 HVGFTF-----PEGFEGVIRKLDYLDKLGITAIEMPTAOPPGKRDGCVGLYAV 161
Db 123 HVGDFSGGDDPARGYKHVIEKLDYLCGLINAIELLPVKEYPGYSWGNPRYFFAT 182
QY 162 QNSYSGPEGRFKLVDEAHKGLGVLVDVYNHVGPEGNY-MVKLGPYFSQKYKTP---WG 217
Db 183 ESSYGSATDLKLVDECHQIRIIMDGIYNHSEASSPLTQIDHDYWHHPRPDNNWG 242
QY 218 LTFNF-----DDAESDEVKFIENVEWIKENYVNDGFRDVAHAIIDTSKHL-----EE 269
Db 243 PFENYHYDENLETYPARKFTGDTVRYWVGVEYHLDGRYDAARQIANYDFMHWIAQBAKK 302
QY 270 IADVVHYNRIVTAESDLNPRVNV--PKEKGCYNIDAQWDDFHSHIAYLTGERQGY 326
Db 303 TAGAKPYN---VAEHPTETSTNLDGPMDCG-----WHSFYHTIRAHICGDT--- 349
QY 327 YTFGNDLDDIVKSYKDVFYVDGKYSNFRKTHGPEVGLDGCNPFVYIQNH-----QV 380
Db 350 -FDLENLKDVI-----DPKRGF-----LGATNVVNYLTNHDHIMVEL 388
QY 381 GNRGKGRIIKLVDRSY-----KIAAALYLLSPYIPMTMGEEYGEENPFYFSDSKL 437
Db 389 GNR-----EIFHDEAFRAKLGTAILMTAVGVPLIWMGEEFYKPK----- 429
QY 438 IQGVREGKKGODTDPQDESTFNASKLSWKIDEEFYSFKILIKMRK 486
Db 430 -----KQDQSLDWTLLGNDLNRSLFDYHKGLIGLRK 462

RESULT 12
Q59242
ID Q59242 PRELIMINARY; PRT; 652 AA.
AC Q59242; 008485;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 1.4-ALPHA-GLUCAN BRANCHING ENZYME (EC 2.4.1.18)
DE (GLYCOPEN BRANCHING ENZYME) (1.4-ALPHA-GLUCAN BRANCHING ENZYME)
DE (AMYLO-(1,4 TO 1,6)TRANSGLUCOSIDASE)
DE (AMYLO-(1,4-1,6)-TRANSGLUCOSYLASE).
GN
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Geobacillus.
OX NCBI_TaxID=1422;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RX MEDLINE=95031021; PubMed=7944355;
RA Takata H., Takaha T., Kuriki T., Okada S., Takagi M., Imanaka T.;
RT "Properties and active center of the thermostable branching enzyme
from Bacillus stearothermophilus.";
RL Appl. Environ. Microbiol. 60:3096-3104(1994).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RA Takata H.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
[3]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RA Takata H., Takaha T., Okada S., Takagi M., Imanaka T.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: FORMATION OF 1,6-GLUCOSIDIC LINKAGES OF
GLYCOPEN.
CC -|- PATHWAY: THIRD STEP IN GLYCOPEN BIOSYNTHESIS.
CC -|- MISCELLANEOUS: OPTIMAL ACTIVITY AT APPROXIMATELY 39 DEGREES
CELSIUS.
CC -|- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
KNOWN AS THE ALPHA-AMYLASE FAMILY.
```

```
DR EMBL; D87026; BAA19588.1; -.
DR InterPro; IPR000461; Alpha-amylase.
DR InterPro; IPR000461; Isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Glycogen biosynthesis; Transference; Glycosyltransferase.
FT ACT_SITE 308 308 BY SIMILARITY.
FT ACT_SITE 351 351 BY SIMILARITY.
FT ACT_SITE 419 419 BY SIMILARITY.
FT ACT_SITE 419 419 BY SIMILARITY.
SQ SEQUENCE 652 AA; 76793 MW; 4591EB414A0E3FEF CRC64;

Query Match 13.4%; Score 399; DB 2; Length 652;
Best Local Similarity 25.2%; Pred. No. 3e-17;
Matches 149; Conservative 96; Mismatches 189; Indels 158; Gaps 31;

QY 4 AYKIDGNEVI-----FTLNAPYQSKSVKL-----KVEKGLYEMER-DEKGYFTITLN 49
Db 22 SYELFGAHVINEGKVGTRFCVWAPHAREVRLVSGFNDWDGTDRLKVKVDEGVMTIVVP 81
QY 50 NVKVRDRYKY--VLDDASEI--PDPASRYQPEGVHGSPSIIQESK--EFNNEFTLKK--- 100
Db 82 ENLEGHLYKYEIVTPDQGVLFKADPAFYSELPRHTAS-IAYDLKGVQWMDQSWKKRRR 140
QY 101 -----EDLIIEIHVGTF-TPEGTF-----EGVIRKLDYLDKLGITAIEMPTAOPPGKRD 150
Db 141 KRIYDQPMVIYELHFGSKKKDKGRFYTYREMADELISYVLDHGFTHELLPLVEHPDLRS 200
QY 151 WGTGCVLYAVQNSYGGPEGRKLVDEAHKGLGVLVDVYNHV--GPEGNYVMKLGCP-- 206
Db 201 WGYQGTGYAVTSRYGTPHDFMYFVDRCHQAGIGVIMDWVPGHFCKDAHGLYMFDCGPTY 260
QY 207 -YFSOKYKT--PWGLTFNFDDAESDEVKFIENVEWIKENYVNDGFRDVAHAII--- 259
Db 261 EYANEKDRNYVMG-TANF-DLGKPEVRSFLISNALFWLEYIHDGFRVDAVANMLYWP 318
QY 260 -----DTSKPHILEIADVVHRY--NRIVIAESDLNDRPVNPKKCGYNIDAQWVDDF 311
Db 319 NDRLYENPYAVEFLRKLNEAVFAYDPNALMAEDSTDWPKVTAPTYEGGLGFYKWK--- 374
QY 312 HHSIAYLTGERGYTDFGNLDDIVKSYKDVFYVDGKYSNFRKTHGPEVGLDGCNFF 371
Db 375 -----NMGMNDMLK-YMETPPYE-----RRHVHNVQVTFSL---LY 406
QY 372 VYIQN-----HDQVGNRGKERIILVDRESYKIAAALYLLSPYI-----PKIEMGEE 419
Db 407 AYSENFILPFSHDEVVH-GKKSLLNKMPGSYEERF-AQLRLLYGYMMAHFGKLLFMGNE 464
QY 420 YGEENPFYFSDFSKLIQGVREGKKGKENGQDTPQDESTFNASKLSWKIDBEI----- 474
Db 465 -----FAQFDE-----WKFEDLDWVLF 482
QY 475 -FSFYKILIKMRKELSTACDRRVNV-----NGENWLLIKREYFSLVVF 518
Db 483 DFLHRKMNDYMKEL-IACYKRYKPFYELDHPDQGFWDHVAEQ-SIFS 532

RESULT 13
Q93Q35
ID Q93Q35 PRELIMINARY; PRT; 440 AA.
AC Q93Q35;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BRANCHING ENZYME GLGB (FRAGMENT).
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN
RP SEQUENCE FROM N.A.
RA Ueki T., Inouye S.;
RT "Identification of a new His-Asp phosphorelay signal transduction
```

RT system which regulates expression of a heat shock gene, lonD, of
RT Myxococcus xanthus.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF285783; AAK83002.1; -
FT NON_TER 440 440
SQ SEQUENCE 440 AA; 48964 MW; 8D492EA5A4A92017 CRC64;

Query Match 13.3%; Score 397; DB 2; Length 440;
Best Local Similarity 28.0%; Pred. No. 2.3e-17;
Matches 131; Conservative 75; Mismatches 164; Indels 98; Gaps 23;

QY 14 FTLMAPYQKSVKLVLEKLY-----EMERDEKGYFTITLNNVVRDRYKYL----- 61
DB 12 FRWAP-----MASRVFVSDFNGWTWELGNEFNFGSGDVAGAVKQGYKFTIRNQWG 67
QY 62 DDASEIPDPASRYQEGVHGSPQIIQESKEF--NNETFLKK--EDLIIEIHVGTG--TP 115
DB 68 SDAWR-ADPSAWQENST--GSSIIYDHGEYWNAAQQYSSPGCFNEWIIYEMHVGTFHDS 124
QY 116 ---EGTFEGVIRKLDGLGTAIEIMPIAOPGKRWGVDGVLVYAVQNSYGGPEGFR 172
DB 125 GFGPGNWSAIALRDLRGANMIKMPAYEFAGDFSGWYNAAPFAPESAYGHPNDMK 184
QY 173 KLVDEAHKKGVLVDVYNHVG-----EGNVMVKLGPFYSQYK--TPWGLTFNF 222
DB 185 REVDEAHMRGIVGIFDVHNNHVGPSDLPMWCFSGDCLSGSGGGEFYFNDRKSTPWGDT--R 242
QY 223 DDAESDEVKRFLENVYWKYNNVDGFRDLDAVHAI-----IDTSPKHILEIAADVHKY 277
DB 243 PDYGRPEVAYIRDSMMNLTHSFRDGLRWDAKYMTQNGSDT---AIPDAWVFRSI 299
QY 278 NR-----TVIAESDLNDPRVYN---PREKCGYNIDAQWDDFHSHIAYLTGERQG 325
DB 300 NREINATQPKISIAEDFGGDFITNDATSDTSGGAGFDSQWGGDFVHAIRAAVIAS--- 356
QY 326 YVDFGNLDDIVKSKVDFV--YDKYSNFRKTHGEPVGLDGCNFVYVYQNHQVGNR 383
DB 357 --NDSGR---DNMSVRNATQIRYSGRHT-----ARVYSESHDEVAN- 393
QY 384 KGRERIKLV-----DREYKIAAALYLLSPYIPMFMEYGE 422
DB 394 GKA-RVPEIWPNGNAGSWAKKRSLAAGVVFSTPGIPIFGQGFLE 440

RESULT 14
Q930Z9 PRELIMINARY; PRT; 1142 AA.
AC Q930Z9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALKALINE PULLULANASE.
GN PUL1876.
OS Bacillus sp. KSM-1876.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=72413;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=KSM-1876;
RA Hatada Y., Ico S.;
RT "alkaline pullulanase from the alkaliphilic bacterium."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049812; BAB47586.1; -
SQ SEQUENCE 1142 AA; 128759 MW; 318FC01362961E84 CRC64;

Query Match 13.3%; Score 395.5; DB 2; Length 1142;
Best Local Similarity 22.9%; Pred. No. 1.1e-16;
Matches 164; Conservative 100; Mismatches 190; Indels 261; Gaps 34;
QY 8 DGNEVITFLWAP-----YQK-----SVKLVLEKGLYEMERDE-----KG 42

DB 376 DGTATI-KLWSPKAERHVQVILYDKNDPDDIVTEVEVMKLGRCGVWEVQLTKEKNTGLDLSRG 434
QY 43 YFTITLNNVVRDRYKYLVD--DASEIPDPASR-----YQPEGVHGSPQIIQ 87
DB 435 YV-----YHYEIDHGDGKRIALDPYAKSLSAWSNDEQGPYAKAALVDPSSIGP 482
QY 88 ESKEFNNETFLKKEDLIIEIHVGTFTPE-----GTTFEGVIRKLDYKLDGITA 136
DB 483 ELEFAHIEGFEKEDIIIEVHVRFDTSDPHIADeltaTQAQGFASFVKLDYIEDLGYTH 542
QY 137 IEIMPIAOF-----PKRD-----WGVDGYLVAVQNSYG-----GPE---G 170
DB 543 IQLLVMSYFFANEFKNDERMLDFSSTNTNMGYDPSQSYFALTGMYSEDTDPDLRIKE 602
QY 171 FRKLVDEAHKKGVLVDVYNHVGPEGNVMVKLGPFYSQYKTPWGLTFNFDDAES--- 227
DB 603 PKKLIDEIHSRGMGVVLVDVYNHRTARVIGIFEDLVNY-----YHFMADAGTPR 650
QY 228 -----DEVKRFLENVYWKYNNVDGFRDLDAVHAIIDTSPKHILEIAADVH 275
DB 651 TSFGGRLGTTHEMSRRILVDSITVWVEEFKIDGFRPDM-----GDHDAESIQAID 703
QY 276 KY-----NRVIAE-----SDLNDPRVYNPKKCGYNIDAQWV-----DDFHHS 314
DB 704 KAKEINPNVIMIGEWIITAGDEDDPNV-----QAADQHMVQYTESVGVFSDEFERNE 755
QY 315 IHA-----YLTGERQGYTDFGNLD--DIVKSKVDFVYDKYNSFRKTHGEPV 362
DB 756 LKSGFGEHGEPRFLTGGAR-----NIDLFDNITKAQPHNFIAD----- 793
QY 363 GELDCNFVYVYQNHQVGNRGKGERIKLVDR-----ESYKIAAALYLLSPYIP 412
DB 794 --DPGDVVPYIEAHD---NLTLHDVIAWAIQKDDHQQEIHQIRLGNLMTLSOGIA 847
QY 413 MIFNGEEV-----EEN--PF---YFFSDFSCKLQGVREGKK 447
DB 848 FLHAGOEYGRTKQFRAETSEPPYKSTYMTDENGEPFRYPYFIHDSYDSTDIINRFDWERA 907
QY 448 ENGODTPODEST-----FNASKLSWK--IDEEIFSYFKILIKMRKELSTAC 492
DB 908 TNA-DATPIQNLREYTTGLIELRRSSDAFLGTHKDLVDEKV---TQNLNPEIETDLV 963
QY 493 DRRVNVNGENWLIKREYFSLVYFSKSSIEVYKSTGLTSSNNSFPQHIEEGK 347
DB 964 AYRIEATGTE-----AFYVFNADDEER---TLTEEDLTVGEFVVDCK 1004

RESULT 15
Q59319 PRELIMINARY; PRT; 825 AA.
AC Q59319;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-DEXTRIN 6-GLUCANOHYDROLASE (EC 3.2.1.41).
GN PULIA.
OS Caldcellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Thermoanaerobacter group; Caldicellulosiruptor.
OX NCBI_TaxID=44001;
RN (1)
RP SEQUENCE FROM N.A.
RA Albertson G.D., McHale R., Gibbs M.D., Bergquist P.L.;
RT "Cloning and expression of a type II pullulanase from an extremely
RT thermophilic anaerobic bacterium, Caldicellulosiruptor
RT saccharolyticus.";
RL Eur. J. Biochem. 0:0-0(1995).
RN (2)
RP SEQUENCE FROM N.A.
RA Gibbs M.D.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; L39876; AAB06264.1; -

Search completed: July 15, 2002, 12:24:08
Job time: 390 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:17:32 ; Search time 58.93 Seconds
(without alignments)
1047.973 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSGGNIENKGIKFLWAP.....KLKDELIKVNRGVGVQLE 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
A_Geneseq_032802.*			
1:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Match Length	Description
1	2954	100.0	556 17 AAR92755 Trehalose-releasin
2	2954	100.0	556 17 AAR90620 Sulfolobus acidoca
3	1748.5	59.2	559 17 AAR90619 Sulfolobus solfata
4	1042.5	35.3	595 22 AAG92072 C glutamicum prote
5	1042.5	35.3	610 22 AAB80206 Corynebacterium gl
6	969	32.8	589 16 AAR80289 Trehalose releasin
7	969	32.8	596 16 AAR77470 Trehalose releasin
8	959	32.5	598 16 AAR80290 Trehalose releasin
9	958	32.4	597 16 AAR85157 Trehalose-releasin
10	936	31.7	575 21 AAY85165 Trehalose-releasin
11	936	31.7	575 21 AAY85165 Trehalose-releasin

12	360	12.2	652	16	AAR00037
13	360	12.2	652	17	AAR96109
14	356	12.1	658	13	AAR23787
15	342.5	11.6	639	12	AAR11271
16	336	11.4	731	22	AAG91100
17	336	11.4	731	22	AAB79423
18	326.5	11.1	648	10	AAP94635
19	323	10.9	772	18	AAW34567
20	321	10.9	772	19	AAW49871
21	320	10.8	793	20	AAW00869
22	313.5	10.6	1250	21	AAW91279
23	312	10.6	670	22	AAU60913
24	310.5	10.5	606	18	AAW36602
25	306	10.4	718	22	AAU60903
26	306	10.4	766	20	AAU00870
27	303.5	10.3	666	20	AAW34991
28	301	10.2	562	10	AAW91904
29	298.5	10.1	621	21	AAW19285
30	296.5	10.0	818	20	AAW17523
31	296	10.0	630	22	AAW69074
32	293.5	9.9	931	20	AAW27357
33	292	9.9	829	22	AAE05691
34	292	9.9	862	21	AAW78513
35	285	9.6	764	21	AAW50819
36	284	9.6	798	20	AAW73552
37	284	9.6	1938	20	AAW73553
38	282	9.5	893	17	AAW09257
39	282	9.5	1938	17	AAW09255
40	280.5	9.5	931	20	AAW00872
41	278.5	9.4	618	20	AAW37184
42	277	9.4	921	22	AAE05723
43	277	9.4	921	22	AAE05726
44	276	9.3	762	21	AAW90977
45	276	9.3	921	22	AAE05689

ALIGNMENTS

RESULT 1	
AAR92755	
ID	AAR92755 standard; Protein; 556 AA.
XX	
AC	AAR92755;
XX	
DT	03-AUG-1996 (first entry)
XX	
DE	Trehalose-releasing thermostable enzyme.
XX	
KW	Thermostable enzyme; trehalose; sweetener.
XX	
OS	Sulfolobus acidocaldarius strain ATCC 33909.
XX	
PN	AU9527131-A.
XX	
PD	01-FEB-1996.
XX	
PF	21-JUL-1995; 95AU-0027131.
XX	
PR	04-JUL-1995; 95JP-0189760.
PR	21-JUL-1994; 94JP-0190180.
PR	11-APR-1995; 95JP-0109128.
XX	
PA	(HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX	
PI	Kubota M, Mitsuami H, Sugimoto T;
XX	
DR	WPI; 1996-106284/12.
DR	N-PSDB; AAT16899.
XX	
PT	Recombinant thermostable enzyme from Sulfolobus acidocaldarius, -
PT	releases trehalose from non-reducing saccharide at temps. exceeding
PT	55 degrees Centigrade

```
XX PS Claim 2; Page 53-54; 74pp; English.
XX PS
XX CC A thermostable enzyme (AAR92755) of Sulfolobus acidocaldarius ATCC
CC 33909 releases trehalose from non-reducing saccharides having a
CC trehalose structure as an end unit and a degree of polymerisation of
CC at least 3. It has a mol.wt. of 54-64 kDa (SDS-PAGE), a pI of
CC 5.6-6.6 and is substantially not inactivated when incubated in aq.
CC solution (pH 7.0) at 85 deg for 60 min. Recombinant enzyme is
CC obtd. by expression of an isolated DNA fragment (AAT16899) in
CC host cells, pref. Escherichia coli, using e.g. vector pBluescript
CC II SK(+). The trehalose is useful as a sweetener.
XX SQ Sequence 556 AA;

Query Match 100.0%; Score 2954; DB 17; Length 556;
Best Local Similarity 100.0%; Pred. No. 7.5e-267;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSFGGNIERNKGIFKLWAPYVNSVKLSKLLIPMEKNDEGFEVEIDDEENLTYSYI 60
Db 1 mfsfggniekngifklwapyvnsvklskllipmekndegffevelddeenltsysi 60

Qy 61 IEDKREIPDPASRYOPLGVHDKSQLIRTDYQILDGLGVKIEDLIYELHVGTFQSGNFK 120
Db 1 iedkreipdpasryoplglvhdksqlirtdyqildglgvkiedliyelhvgtfsggnfk 120

Qy 61 GVIEKLDYLKDLGTIGIELMPVAQFPGNRDWDYGVFLYAVQNTYGGPWELAKLVNEAHK 180
Db 61 gviekldylkdlgtigielmpvaqfpgnrwdygvflyavqntyggpwlaklvneahk 180

Qy 181 RGIAVILDVYVNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDGRGCDQVRKFILENVEY 240
Db 181 rgiavildevynhigpegnyllglgpyfsdryktpwgltnfdgrgcdqvrkfilenvey 240

Qy 241 WFKTFKIDGLRLDAVHAIFDPSNPKHILQETAIAEKAHQKGFVIAESDNDPKIVKDDCGYK 300
Db 241 wfktdkldglrldavhaifdpsnphkhlqetaiaekahqkgfviaesdndpkivkddcgyk 300

Qy 301 IDAQWDDFHAVHAFITKEKDYIYQDFRIEDIEKTFKDVYVDGKYSRYGRGTHGAPV 360
Db 301 idaqwddfhavhafitkekdyiyqdfriediektfkdvvyvdgkysrygrgthgapv 360

Qy 361 GDLPPRFVVFVFIQNHQDQVGNRNGERLSILTDKTYLMAATLYILSPYIPIFMGEYYE 420
Db 361 gdlpprkfvvfivfiqnhdqvgngngersiltdktylmaatlyilspyiplfmgeeyye 420

Qy 421 TNPFFFSDFSDPVLIKVREGRLKENNQMDPOSEAFKLSKLSWKIDEBVLDDYKQLI 480
Db 421 tnpfffsdfsdpvlkivregrlkennqmdposseafklsklskwkideevlddykqli 480

Qy 481 NIKRYNNCKRVKVRREGNCITLIMEKIGTIIASFDIVINSKITGNLLIGIFPKLIKK 540
Db 481 nikrynnckrvkvrregncitlimekigtiiasfddivinskitgnlligifpklikk 540

Qy 541 DELIKVNRGVGVYQLE 556
Db 541 delikvnrvgvvyqle 556

RESULT 2
AAR90620
ID AAR90620 standard; Protein; 556 AA.
XX AC
XX AC AAR90620;
XX DT
XX DT 29-JUN-1996 (first entry)
XX XX Sulfolobus acidocaldarius amylase for alpha, alpha-trehalose prodn.
DE DE transferase: amylase; Sulfolobus; production; alpha, alpha-trehalose;
XX KW
XX KW malto-oligosaccharide; hydrolysis.
```

```
XX OS Sulfolobus acidocaldarius.
XX PN WO9534642-A.
XX PD 21-DEC-1995.
XX PF 14-JUN-1995; 95WO-JP01189.
XX PR 21-APR-1995; 95JP-0120673.
XX PR 15-JUN-1994; 94JP-0133354.
XX PR 18-AUG-1994; 94JP-0194223.
XX PR 31-OCT-1994; 94JP-0290394.
XX PR 21-NOV-1994; 94JP-0286917.
XX PR 21-NOV-1994; 94JP-0311185.
XX PA (KIRI ) KIRIN BEER KK.
XX WP1; 1996-049671/05.
XX DR N-PSDB; AAT12326.
XX PT Sulfolobus spp. derived transferase and amylase - for production of
XX alpha, alpha-trehalose from malto-oligosaccharide(s)
XX PS Claim 112; Page 244-250; 357pp; Japanese.
XX CC The amylase is derived from Sulfolobus acidocaldarius. The amylase acts
CC on a saccharide having at least three sugar units, which are pref.
CC glucose units at the reducing end (the linkage between the first and
CC second glucose units is alpha-1, alpha-1, while the linkage between the
CC second and third glucose units is alpha-1,4), to hydrolyse alpha-1,4
CC linkages within the sugar chain, yielding alpha, alpha-trehalose and also
CC mono- and disaccharide hydrolysis products. The amylase has a mol. wt. of
CC 61 to 64 kDa. It is characterised by working at pH 4.5-5.5 and at 60-85
CC deg.C. It has an isoelectric point of 4.3-5.4 and retains at least
CC 100 percent activity after 6 hrs. at 80 deg.C. It is completely inhibited
CC by 5 mM copper sulphate. Use of a transferase and the amylase in
CC succession on suitable substrates such as malto-oligosaccharides, is
CC useful for the production of alpha, alpha-trehalose.
XX SQ Sequence 556 AA;

Query Match 100.0%; Score 2954; DB 17; Length 556;
Best Local Similarity 100.0%; Pred. No. 7.5e-267;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSFGGNIERNKGIFKLWAPYVNSVKLSKLLIPMEKNDEGFEVEIDDEENLTYSYI 60
Db 1 mfsfggniekngifklwapyvnsvklskllipmekndegffevelddeenltsysi 60

Qy 61 IEDKREIPDPASRYOPLGVHDKSQLIRTDYQILDGLGVKIEDLIYELHVGTFQSGNFK 120
Db 61 iedkreipdpasryoplglvhdksqlirtdyqildglgvkiedliyelhvgtfsggnfk 120

Qy 121 GVIEKLDYLKDLGTIGIELMPVAQFPGNRDWDYGVFLYAVQNTYGGPWELAKLVNEAHK 180
Db 121 gviekldylkdlgtigielmpvaqfpgnrwdygvflyavqntyggpwlaklvneahk 180

Qy 181 RGIAVILDVYVNHIGPEGNYLLGLGPYFSDRYKTPWGLTFNFDGRGCDQVRKFILENVEY 240
Db 181 rgiavildevynhigpegnyllglgpyfsdryktpwgltnfdgrgcdqvrkfilenvey 240

Qy 241 WFKTFKIDGLRLDAVHAIFDPSNPKHILQETAIAEKAHQKGFVIAESDNDPKIVKDDCGYK 300
Db 241 wfktdkldglrldavhaifdpsnphkhlqetaiaekahqkgfviaesdndpkivkddcgyk 300

Qy 301 IDAQWDDFHAVHAFITKEKDYIYQDFRIEDIEKTFKDVYVDGKYSRYGRGTHGAPV 360
Db 301 idaqwddfhavhafitkekdyiyqdfriediektfkdvvyvdgkysrygrgthgapv 360

Qy 361 GDLPPRFVVFVFIQNHQDQVGNRNGERLSILTDKTYLMAATLYILSPYIPIFMGEYYE 420
Db 361 gdlpprkfvvfivfiqnhdqvgngngersiltdktylmaatlyilspyiplfmgeeyye 420
```


Db 361 gdlpprkfvvfiqhndqvgngerslsldtktylmaatlyllspylplifmggeeye 420
QY 421 TNPFFFSDFSDPVLKIGVREGRLKENNOMIDPQSEAFKLSKLSWKIDBEVLDYKQLI 480
Db 421 tnpfffsdfsdplkigvregrlkennqmidpqqseafklsklskwideevldyvkli 480
QY 481 NIKRYNNCKRVKEVRREGNCITLIMEKIGITIASFDODIVINSKITGNLLIGIFPKKLKK 540
Db 481 nlrkrynnckrvkevrregncitlimekigilasfddivinskitgnlligifpkklkk 540
QY 541 DELIKVNRGVGVYQLE 556
Db 541 delikvnrvgvvyqle 556

RESULT 3
AAR90619
ID AAR90619 standard; Protein; 559 AA.
AC AAR90619;
XX
DT 29-JUN-1996 (first entry)
XX
DE Sulfolobus solfataricus amylase for alpha, alpha-trehalose prodn.
XX
KW transferase; amylase; Sulfolobus; production; alpha, alpha-trehalose;
KW malto-oligosaccharide; hydrolysis.
XX
OS Sulfolobus solfataricus.
XX
PN WO9534642-A.
XX
PD 21-DEC-1995.
XX
PF 14-JUN-1995; 95WO-JP01189.
XX
PR 21-APR-1995; 95JP-0120673.
PR 15-JUN-1994; 94JP-0133354.
PR 18-AUG-1994; 94JP-0194223.
PR 31-OCT-1994; 94JP-0290394.
PR 21-NOV-1994; 94JP-0286917.
PR 21-NOV-1994; 94JP-0311185.
XX
PA (KIRI) KIRIN BEER KK.
XX
DR WPI; 1996-049671/05.
DR N-PSDB; AAT12325.
XX
PT Sulfolobus spp. derived transferase and amylase - for production of
PT alpha, alpha-trehalose from malto-oligosaccharide(s)
XX
PS Claim 108; Page 235-240; 357pp; Japanese.
XX
CC The amylase is derived from Sulfolobus solfataricus. The amylase acts
CC on a saccharide having at least three sugar units, which are pref.
CC glucose units at the reducing end (the linkage between the first and
CC second glucose units is alpha-1, alpha-1, while the linkage between the
CC second and third glucose units is alpha-1,4), to hydrolyse alpha-1,4
CC linkages within the sugar chain, yielding alpha, alpha-trehalose and also
CC mono- and disaccharide hydrolysis products. The amylase has a mol. wt. of
CC 61 to 64 kDa. It is characterised by working at pH 4.5-5.5 and at 60-85
CC deg.C. It has an isoelectric point of 4.3-5.4 and retains at least
CC 100 percent activity after 6 hrs. at 80 deg.C. It is completely inhibited
CC by 5 mM copper sulphate. Use of a transferase and the amylase in
CC succession on suitable substrates such as malto-oligosaccharides, is
CC useful for the production of alpha, alpha-trehalose.
XX
SQ Sequence 559 AA;

Query Match 59.2%; Score 1748.5; DB 17; Length 559;
Best Local Similarity 58.9%; Pred. No. 2.9e-154;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;

QY 3 SFGGNIEKNKGIKFLWAPVNSVYKVK-LSKKLIPMEKNDGEFFVEIDIEENLTYSYII 61
Db 2 tfaykidgnevifltlwapyqskvklvlekglyemerdekgyftitlnnvkvrdrkyvl 61
QY 62 EDKREIPDPASRYPOLGVHDKSQILRTDYLGLGVKVIETDLIIYELHVGTFQSGNFKG 121
Db 62 ddaseipdpasryqpegvhgpgsqliqeskefnnetfikkedlliyehvgtftpegtfeg 121
QY 122 VIEKLDYLDLGTITGIELMPVAQPGNRDMGYDGVFLYAVONTYGGPFWELAKLVNBAHKK 181
Db 122 virklayldlgtitaieimpiaqfpgkrdwgydgvlyavqnsyggpegfrklvdeahkk 181
QY 182 GIAVILDVYVNHIGPEGNYLLGLPYFSDRYKTPWGLTFNFDRCQDVVKFLENVYEW 241
Db 182 glgvildvynhvpgpegymvklgpyfsqkyktpwgltfnfdaesdevrkfilenveyw 241
QY 242 FKTFKIDGLRLDAVHAIFDNPSPKHILOETAEKAOHLGKFVIAESDLNDPKIV--KDDCGY 299
Db 242 ikeynvdgfrldavhaidtspkhileetadvvhkynriviaesdlndprvnpkckgy 301
QY 300 KIDAQWDDFHHAVHAFITKEKDYIYQDFGRIEDIEKTKDFVYDGYKSYRGTRTHGAP 359
Db 302 nidaqwwddfhshihayltgerggytdfgnlddiwksykdvfvydgkysnfrkrthgep 361
QY 360 VGLPFRPKFVVFIONHDQVGNRNGRERLSLTIDTKTYLMAATLYILSPYIPLIFMGEEYY 419
Db 362 vgelgcnfvvyiqnhdqvgngrgkeriiklvdresyktaaalylispylpmifmggeyg 421
QY 420 ETNPFFFSDFSDPVLKIGVREGRLKENNOMIDPQSEAEFLKLSWKIDBEVLDYKQL 479
Db 422 eepfyfssdfsdkllqgvregkrkngqtdpqdestfnasklskwkideeifsfykil 481
QY 480 INIRKRYN-NCKRVKEVRREGNCITLIMEKIGITIASFDDIVINSKITGNLLIGI--GFPK 536
Db 482 ikmrkelsiactdrvnvngengenwiikgreysiyvfkssievkysgtllissnnsfpq 541
QY 537 KLKKDELIKVNRGVGVYQL 555
Db 542 hieegk-yefdkgfalykl 559

RESULT 4
AAG92072
ID AAG92072 standard; Protein; 595 AA.
XX
AC AAG92072;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 5826.
XX
KW Corynebacterium glutamicum.
XX
OS Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX
PN Corynebacterium glutamicum.
XX
PD BP1108790-A2.
XX
PF 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.

PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042124.
PR 03-SEP-1999; 99DE-1042129.
PR 09-MAR-2000; 2000US-0187970.
XX (BADI) BASF AG.
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX WPI; 2001-137957/14.
DR N-PSDB; AAF72325.
XX Nucleic acids from Corynebacterium glutamicum encoding metabolic
PT pathway proteins, useful for producing fine chemicals in
PT microorganisms, including organic acids, nonproteinogenic amino acids,
PT and purine and pyrimidine bases -
XX
PS Claim 20; Page 1722-1725; 1737pp; English.
XX AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211: The C. glutamicum
CC MP nucleic acids are useful for the production of fine chemicals
CC in microorganisms, including organic acids, nonproteinogenic amino
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic
CC compounds, vitamins, cofactors, polyketides and enzymes.
XX
SQ Sequence 610 AA;

Query Match 35.3%; Score 1042.5; DB 22; Length 610;
Best Local Similarity 40.8%; Pred. No. 2.8e-88;
Matches 225; Conservative 93; Mismatches 199; Indels 35; Gaps 13;

QY 15 FKLWAPYVNSVKLKSLLIPMEKDEGFVEIDIEENLTYSYIIE----KREIPDP 70
DB 43 fswaplphdvhlilngeltpmhktgswraeiap-kagdrygfslfdgsswsktlpdp 101
QY 71 ASRYQLGVHDKSOLRTYQIILD---LGKVKIEDLIIYELHVGTFQSGNFKGVIEKLD 127
DB 102 rstqpgdgvhlsevsddysldgqwtgri-lpgsvlyelhvgtfsgdtfgvvdtkip 160
QY 128 YLKDLGITGLMPLVAFQFPGNDRMGVDFLYAVQNTYGPWELAKLVNEAKRGIAVIL 187
DB 161 Yirdlgtvtaellpvqfpggnrnwgydglwlvhavhagyg9pgaglkklidashqaglavyl 220
QY 188 DVYNNHIGPEGNYLLGLGYPFSRYKTPWGLTFNFDRCGDQVRKFTLENVEYWFRTFKI 247
DB 221 dvynnhfpgdngnvgfpytsag-stgwgdvvnvngdhdsdevrnylldaarqwfedfhv 279
QY 248 DGLRLDAVHAIFNPSKPHILOETAEAKHQLG-----KVFIAESDLNDPKIV--KDCGY 299
DB 280 dgirldavhslldrgayslaqltmvaedvsdqgiprsliaeslndpkfvtsreaggf 339
QY 300 KIDAQWDDFHAVHAFITKEKDYDYODFGRIEDIEKTFKDVYDGYKSYRGRGTHGAP 359
DB 340 gidaqvwddhhahlvsgerngyysdfgsdvdtlaktirevfchngystygrnhgtp 399
QY 360 VGD--LPPRKVFQIHHQVQNGRNGERLS-ILTDKTYLMAATYILSPYIPLIFMGE 416
DB 400 vhpdpasarfvtthtdqtnraigdrpstltltpceqqvkaaiiy-sspytpmlfmg 458
QY 417 EYETNPFFFSDFSPVLKGVREGRKLE-----NNQIMDPQSEEAFLSKLSWKID 469
DB 459 efgattpaffcshtdpelnrltsegrkrefarlgwnaddipspelestfsskdwft 518
QY 470 EE---VLDYKQLINIRKRYNCK---RVKEVRENGCITLIMEKIGITASFDDIVNSK 523
DB 519 aeqrindaykqlilhrhtlfgsqpnlltlevehgenwismangrgrilansddtitvp 578

QY 524 ITGNLLIGIGFP 535
DB 579 Iggellysftsp 590

RESULT 6
ID AAR80289 standard; Protein; 589 AA.
XX AAR80289;
XX 19-JAN-1996 (first entry)
XX Trehalose releasing enzyme.
XX Enzyme: trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW alpha-maltotriosyltrehalose; alpha-maltotetraosyltrehalose;
KW maltopentaosyltrehalose; sweeter; taste-improving agent;
KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW pharmaceuticals.
XX Rhizobium sp. M11.
XX EP671470-A2.
XX 13-SEP-1995.
XX 07-MAR-1995; 95EP-0301474.
XX 07-MAR-1994; 94JP-0059840.
XX 07-MAR-1994; 94JP-0059834.
XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX WPI; 1995-312772/41.
XX N-PSDB; AAQ98669.
XX DNA encoding a trehalose releasing enzyme - which releases trehalose
PT from a non-reducing saccharide having a trehalose structure as an
PT end unit.
XX Claim 3; Page 21-22; 45pp; English.
XX This enzyme can be used for the preparation of trehalose with high
CC yields and efficiency from non-reducing saccharides such as
CC alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC alpha-maltotriosyltrehalose, alpha-maltotetraosyltrehalose and
CC maltopentaosyltrehalose. The trehalose can be used as a sweetener,
CC taste-improving agent, quality-improving agent, stabiliser, filler,
CC excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX Sequence 589 AA;

Query Match 32.8%; Score 969; DB 16; Length 589;
Best Local Similarity 39.5%; Pred. No. 2e-81;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

QY 13 GIFKLWAPYVNSVKLKSLLIPMEKN-----DEGFVEIDIEENLTYSYIIE-DKR 65
DB 8 grfdwaapeagtvtllageryemgrrpgngpadegwtaadaptdgadvgyllgdgel 67
QY 66 ETPDPSRQPLGVHDKSOLRTYQIILDGLKVKIEDL-----IYELHVGTFQ 115
DB 68 pldprrtrrqqegvh---alsrt----fdpgahrwdagwqgrglgsgviyelhigtft 120
QY 116 EGNFKGVIEKLDYLDLGTGIELMPVAFQFPGNDRMGVDFLYAVQNTYGPWELAKLV 175
DB 121 egtldaaagklidylagldgfellpvnafngthnwgdygqvfvahnegy9gpaaygrfv 180
QY 176 NEAKRGIAVILDVYNNHIGPEGNYLLGLGYPFSRYKTPWGLTFNFDRCGDQVRKFI 235

Db 181 daahaaglvigdvvynhlgpsnylprygyllkhgegnwgdsvnlodgpgsdhvrqyl 240
Qy 236 ENVEYFTEKIDGLRLDAVHAIFDNPSPKHILQETIAEKAHOLGK-----FVIAESDLND 289
Db 241 dnvamwlrdrvvdglridavhalkderavhileefgaladalseseggrptlilaesdlnd 300
Qy 290 PKIV--KDDCGYKIDAQWDDFHHAHAFITKEKDYQDFGRIDIEKTFKDVVYDVK 347
Db 301 prllprdvngyglagqwsddfhbhavhvnvsgettgyysdfdsigalakvldgffhdgs 360
Qy 348 YSRVGRTHGAPV--GDLPPRKFFVFTQNDQVGNRNGERLSLTLDKTYLMAATLYL 405
Db 361 yssfrgrchprlnfsavhpaalvvcsgnhdqignratgdrslsglpygsalaaavltlt 420
Qy 406 SPYIPLFMGEYETNPFFSDFSDPVLKGVREGRLKENNOM-----IDPQSEEA 458
Db 421 gpftplmfmeegygattpwffftshpepelgkataegrriferemgdwpavvpdpqpet 480
Qy 459 FLKSLSWK-----IDEEVDYKQLINIRKRYNNCKRVKVRREGNCITLIMEKIGIIA 513
Db 481 ftrskldwaeasagdhalllyrslitlr-----rstpelarlgtadv----- 526
Qy 514 SFDD 517
Db 527 efdd 530
RESULT 7
AAR77470
ID AAR77470 standard; Protein: 596 AA.
XX
AC AAR77470;
XX
DT 25-JAN-1996 (first entry)
XX
DE Trehalose releasing enzyme.
XX
KW Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW alpha-maltotriosyltrehalose; alpha-maltotetraosyltrehalose;
KW maltopentaosyltrehalose; sweetener; taste-improving agent;
KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW pharmaceuticals.
XX
OS Rhizobium sp. M11.
XX
PN EP671470-A2.
XX
PD 13-SEP-1995.
XX
PF 07-MAR-1995; 95EP-0301474.
XX
PR 07-MAR-1994; 94JP-0059840.
PR 07-MAR-1994; 94JP-0059834.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
PI Hattori K, Kubota M, Sugimoto T, Tsusaki K;
XX
DR WPI: 1995-312772/41.
DR N-PSDB; AAQ98671.
XX
PT DNA encoding a trehalose releasing enzyme - which releases trehalose
PT from a non-reducing saccharide having a trehalose structure as an
PT end unit.
XX
PS Claim 3; Page 27-29; 45pp; English.
XX
CC This enzyme can be used for the preparation of trehalose with high
CC yields and efficiency from non-reducing saccharides such as
CC alpha-glucosyltrehalose, alpha-maltosyltrehalose,
CC alpha-maltotriosyltrehalose, alpha-maltotetraosyltrehalose and
CC maltopentaosyltrehalose. The trehalose can be used as a sweetener,

CC taste-improving agent, quality-improving agent, stabiliser, filler,
CC excipient or adjuvant in food products cosmetics and pharmaceuticals.
XX
SQ Sequence 596 AA;
Query Match 32.8%; Score 969; DB 16; Length 596;
Best Local Similarity 39.5%; Pred. No. 2e-81;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;
Qy 13 GIFKLWAPVNSVSKLSKLIPEKN-----DEGFEVEIDIEENLYSYLIE-DKR 65
Db 15 grldwapeagvtvllageryemgrrpgnpgadegwtaadaptdadvdygylldgdei 74
Qy 66 EIPDPASRYOPLGVHDKSQLIRTDYQILDGLKVKIEDL-----IIVELHVGTFQ 115
Db 75 plpdrtrrpegvh---alsrt-----fdpgahwqdagwqgrlgelqsvlyelhigftfp 127
Qy 116 EGNFKGVIEKLDYLDKDLGITGIELMPVAQFPNGRDWGYGVFLYAVQNTYGGPWELAKLV 175
Db 128 egtldaaagkldylagldfiedfllpvnafngthnwgdygvqwfavhegyggaaygrfv 187
Qy 176 NEAHKRGIAVILDVYVNHIGPEGNYLLGLGPYFSDRYKTPMGLTFNFDRCDDQVRKFI 235
Db 188 daahaaglvigdvvynhlgpsnylprygyllkhgegnwgdsvnlodgpgsdhvrqyl 247
Qy 236 ENVEYFTEKIDGLRLDAVHAIFDNPSPKHILQETIAEKAHOLGK-----FVIAESDLND 289
Db 248 dnvamwlrdrvvdglridavhalkderavhileefgaladalseseggrptlilaesdlnd 307
Qy 290 PKIV--KDDCGYKIDAQWDDFHHAHAFITKEKDYQDFGRIDIEKTFKDVVYDVK 347
Db 308 prllprdvngyglagqwsddfhbhavhvnvsgettgyysdfdsigalakvldgffhdgs 367
Qy 348 YSRVGRTHGAPV--GDLPPRKFFVFTQNDQVGNRNGERLSLTLDKTYLMAATLYL 405
Db 368 yssfrgrchprlnfsavhpaalvvcsgnhdqignratgdrslsglpygsalaaavltlt 427
Qy 406 SPYIPLFMGEYETNPFFSDFSDPVLKGVREGRLKENNOM-----IDPQSEEA 458
Db 428 gpftplmfmeegygattpwffftshpepelgkataegrriferemgdwpavvpdpqpet 487
Qy 459 FLKSLSWK-----IDEEVDYKQLINIRKRYNNCKRVKVRREGNCITLIMEKIGIIA 513
Db 488 ftrskldwaeasagdhalllyrslitlr-----rstpelarlgtadv----- 533
Qy 514 SFDD 517
Db 534 efdd 537
RESULT 8
AAR77471
ID AAR77471 standard; Protein: 598 AA.
XX
AC AAR77471;
XX
DT 25-JAN-1996 (first entry)
XX
DE Trehalose releasing enzyme.
XX
KW Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;
KW alpha-maltotriosyltrehalose; alpha-maltotetraosyltrehalose;
KW maltopentaosyltrehalose; sweetener; taste-improving agent;
KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;
KW pharmaceuticals.
XX
OS Arthrobacter sp. Q36.
XX
PN EP671470-A2.
XX
PD 13-SEP-1995.
XX

Db 317 vngylegqwdhfhavhvtgettygsdfdsaaakavlrdgffhdsgysfrerh 376
Qy 356 HGAPV--GDLPPKFPVFNQNDVNGRNGERLSILTDKTYTLMAATLYILSPYIPLIF 413
Db 377 hgrpinfsavhpaalvcshqndignratdrisqtlpygsalaaavltitgpfpmil 436
Qy 414 MGEYYFNPFPPFSDPVLIRGVREGLKENQM-----IDPQSEAFILKSLWS 466
Db 437 mgeeygastpwwftshpepelgkateagrikefermgwdpavvpdpqdpetfirskldw 496
Qy 467 KIDEE-----VLDYKQLINIRK 484
Db 497 aeaagdhariilelyrsitairr 519

RESULT 10
AAV85157
ID AAV85157 standard; protein; 575 AA.
XX
AC AAV85157;
XX
DT 23-JUN-2000 (first entry)
XX
DE Trehalose-releasing enzyme amino acid sequence.
XX
KW Non-reducing saccharide-forming enzyme; trehalose; disaccharide; food;
KW high yield; trehalose releasing enzyme; moisture retention; cosmetic;
KW production.
XX
OS Arthrobacter sp. S34.
XX
PN EP990704-A2.
XX
PD 05-APR-2000.
XX
PF 13-SEP-1999; 99EP-0307220.
XX
PR 11-SEP-1998; 98JP-0258394.
XX
PR 11-DEC-1998; 98JP-0352252.
XX
PR 26-JAN-1999; 99JP-0016931.
XX
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX
XX Yamamoto T, Maruta K, Kubota M, Fukuda S, Miyake T;
XX WPI: 2000-273225/24.
XX N-PSDB: AAA10503.
XX
XX New mesophilic enzyme from Arthrobacter, used to convert starch
XX hydrolystate into non-reducing sugars, particularly trehalose, useful as
XX a sweetener in foods, pharmaceuticals and cosmetics -
XX
XX Claim 29; Page 47-48; 93pp; English.
XX
XX This sequence represents a trehalose-releasing enzyme amino acid sequence
XX from Arthrobacter sp. S34. This enzyme specifically hydrolyses a
XX non-reducing saccharide having a trehalose structure as an end unit. The
XX invention relates to this enzyme and also to a non-reducing saccharide
XX forming enzyme that forms a non-reducing saccharide having a trehalose
XX structure as an end unit by reducing partial starch hydrolysates. The two
XX enzymes can be used to produce trehalose, a disaccharide consisting of
XX two moles of glucose bound at their reducing residues. The disaccharide
XX is substantially free of reducibility and has moisture retaining
XX abilities. The enzymes are used to produce non-reducing sugars, they are
XX particularly used for further enzymatic conversion of starch hydrolysate
XX to trehalose. The non-reducing sugars produced by the enzymes are useful
XX in foods, pharmaceuticals and cosmetics. Non-reducing sugars are
XX compatible with amino acids and proteins, they do not cause browning, and
XX have good moisture-retaining properties. The enzymes produce the
XX non-reducing sugars in high yield and function at a medium temperature
XX range i.e. between 40 degrees celsius and 60 degrees celsius, and in an
XX acidic pH range, i.e. a pH of less than 7. Production of non-reducing

CC sugars.using the enzymes eliminates the need to convert a starch
CC hydrolystate into sugar alcohols (non-reducing) by catalytic
CC hydrogenation.
XX
SQ Sequence 575 AA;

Query Match 31.7%; Score 936; DB 21; Length 575;
Best Local Similarity 38.7%; Pred. No. 2.3e-78;
Matches 198; Conservative 82; Mismatches 177; Indels 54; Gaps 10;

Qy 15 FKLWAPYVNVSKLKL--SKKLIPMEKNDGFEFE--IDDTEENLTYSIIEDKREIPDP 70
I :||| I :
Db 5 fpvwapaqaqvtlvvgggraelpitrdengwwalqpdpdgdvdygylvdgkpfadp 64
I :||| I :
Qy 71 ASRYQLGVHDKSQLIRTDYQILDGKV-----KIEDLIYELHVGTFNS 114
I :||| I :
Db 65 rslrqprgvh-----elgrfeparyawgdggrgrdltgavlyehlvgtft 111
I :||| I :
Qy 115 QEGNFKGVIKLDYLDIGITGIELMPVAQPPGNRDMGYDGVFLYAVQNTYGGPWELAKL 174
I :||| I :
Db 112 pegtldsairrlhhlvrlgvdaveilpvnafngthgwydgvlwyavhepyggpeayqrf 171
I :||| I :
Qy 175 VNEAHKRGIAVILDVYNHIGPEGNYLLGLQPYFSDRYKTPWGLTFNFDRCDCQVRKFI 234
I :||| I :
Db 172 vdacharglavvqvynhigpsghlpdfgpylgsgaantwgdalnldgplsdevryi 231
I :||| I :
Qy 235 LENVEYFKTFKIDGLRLDAVHAIFDNSPKHILQFIA---EKAHOLGK--FVIAESDLN 288
I :||| I :
Db 232 idnawyrlrdmhgdgrlrdavhralrdaralhlleelaarvdelagegrpltilaesdl 291
I :||| I :
Qy 289 DPKIVKDDC--GYKIDAQWVDDFHVHAFITKEKDYQDFGRIEDIEKTFKDFVFDG 346
I :||| I :
Db 292 dplirsraahgyldaqwdvvhavhantgetvgyadfglgalkvfgqrgwfhg 351
I :||| I :
Qy 347 KYSRYGRTHGAPVG-DLPPRKFFVFIONHDQVNGRNGERLSILTDTKTYLMAATLYIL 405
I :||| I :
Db 352 twssfrerhgrpldpdipfrlrvafaghdqvgnravgdrrmsaqvgegsaaaaalvil 411
I :||| I :
Qy 406 SPYIPLIFMGEYYETNPFFSDFSDPVLIRGVREGLKENQM-----IDPQSEEA 458
I :||| I :
Db 412 gptcmifmgeewgartpwwftshpepelgeatargiaefarmgwdpavvpdpqdp 471
I :||| I :
Qy 459 FLKSKLSWKIDEE-----VLDYKQLINIRK 484
I :||| I :
Db 472 farshldwseperphagllafytdlialrr 502
I :||| I :

RESULT 11
AAV85165
ID AAV85165 standard; protein; 575 AA.
XX
AC AAV85165;
XX
DT 23-JUN-2000 (first entry)
XX
DE Trehalose-releasing enzyme amino acid sequence.
XX
KW Non-reducing saccharide-forming enzyme; trehalose; disaccharide; food;
KW high yield; trehalose releasing enzyme; moisture retention; cosmetic;
KW production.
XX
OS Arthrobacter sp. S34.
XX
PN EP990704-A2.
XX
PD 05-APR-2000.
XX
PF 13-SEP-1999; 99EP-0307220.
XX
PR 11-SEP-1998; 98JP-0258394.
XX
PR 11-DEC-1998; 98JP-0352252.
XX
PR 26-JAN-1999; 99JP-0016931.


```

Db 258 ptveyanekdrenyvwg-tanf-dlgkpevrsflisnalfwleyyhdgfrvdavnmly 315
Qy 256 ----HAIFDNS-PKHILQETAEK--AHQLGKFVIAESDLNDPKIVKD-----DCGYKID 302
Db 316 wpndrlyenpyaveflrklneavfaydpnalmiaedstdwkvptaptgyeglgfnykwn 375
Qy 303 AQWVDDF-----HHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVYDGYKYSRYR 352
Db 376 mgwmndmlkymetppyerhrhngvtfsllays-----enfilpshdevvhgkks--- 427
Qy 353 GRTHGAPVGDLPPrKFVFIQNHQVGNRNG---ERLSILTDKTTVYMAATLYILSPYI 409
Db 428 -----linkmpgsyeekfaqlrllygymma-----hpgk 456
Qy 410 PLIFMGEEYETNPFFFSDFDPVLKGVREGRLKENNOMIDPQSEAEFLSKLSW--- 466
Db 457 klifmgnef-----aqfdewkfedeldwvlf 482

Qy 467 --KIDEVLDYKOLINIRKRY 486
Db 483 dfeihrkmdymkeliacykry 504

```

RESULT 13

AAR96109 ID AAR96109 standard; Protein; 652 AA.

XX AAR96109;

XX 20-AUG-1996 (first entry)

XX Starch-branching-enzyme.

XX Starch-branching-enzyme; thermostable; Q-enzyme; Escherichia coli;
 KW starch; amylopectin; glucan; cyclodextrin; food; food-additive;
 KW adhesive; biodegradable plastic.

XX Bacillus stearothermophilus strain TRBE14 (FERM P-13916).

XX Key Location/Qualifiers

XX Peptide 238..249
 FT /note= "Conserved amylase sequence used to
 FT construct primer AAT27126"

FT Peptide 309..317

FT /note= "Conserved amylase sequence"

FT Peptide 414..423

FT /note= "Conserved amylase sequence used to
 FT construct primer AAT27127"

XX EP710674-A2.

XX 08-MAY-1996.

XX 13-SEP-1995; 95EP-0250222.

XX 31-JUL-1995; 95JP-0195647.

XX 13-SEP-1994; 94JP-0218554.

XX (EZAKI) EZAKI GLICO CO LTD.

XX Fujii K, Imanaka T, Nakamura H, Okada S, Takaha T;

XX Takata H, Terada Y, Yanase M;

XX WPI; 1996-223407/23.

XX N-PSDB; AAT27128.

XX New glucan with cyclic, branched core and outer branched region -
 PT useful in foods, drinks, adhesives, infusion compsns. etc., readily
 PT soluble in water forming solns. of low viscosity

XX Example A-2; Page 28-31; 50pp; English.

XX

CC This sequence represents a thermostable starch-branching-enzyme
 CC (Q-enzyme, EC-2.4.1.18) from Bacillus stearothermophilus TRBE14
 CC (FERM P-13916). Conserved regions present in amylases are used to
 CC construct primers AAT27126-7, which are used in PCR to amplify a
 CC fragment of the gene, which is then used as a probe to isolate a
 CC full-length clone in Escherichia coli. The recombinant enzyme may
 CC be reacted with starch or amylopectin to produce a glucan with a
 CC degree of polymerisation of at least 50, with an inner branched
 CC cyclic region and an outer branched region. The product is freely
 CC soluble in water, forming a low-viscosity solution, is more
 CC resistance to degradation than native starch, does not react with
 CC proteins and amino acids, and is degraded to glucose in vivo, so is
 CC easily digestible. The glucan may be used in foods, beverages,
 CC food-additives, adhesives, in cyclodextrin production and as a
 CC starch substitute in biodegradable plastic.

XX Sequence 652 AA;

Query Match 12.2%; Score 360; DB 17; Length 652;
 Best Local Similarity 25.3%; Pred. No. 1.7e-24;
 Matches 142; Conservative 72; Mismatches 186; Indels 162; Gaps 25;

```

Qy 4 FGNIEKNKG----IFKLWAPYNSVKL-----KLSKKLIPMEKNDGPFPEVE 47
Db 26 fgahvineggkvgrfcwqapharevrlvgvsnfdwdgtdfrlek-----vndegvwtiv 79
Qy 48 IDDEENLITYSY-IIEDKREI---PDPAARYQPLGVHDKSOLIRTD---YQILDLG---- 96
Db 80 vpenleghlykyeivtpdgqvlfkadpyafyselrphstas---laydlkgvqndqswkrk 137
Qy 97 ----KVKIEDLIIVELHVGTF-SQEGNFKGVIEKLD----YKDLGITGITELMPVAFPG 147
Db 138 krrkriyqpmviyelhfgswkkkgdgrfytyremadelisvldhgftthiellplvehpl 197
Qy 148 NRWDYDGVFLYAVONTYGGFWELAKLVNEAHKRGIAVILDVYVNHIT--GPEGNYLLGLG 205
Db 198 drswgvggtgyyavtsrygtphdfmyfvrchqdegigvmdwvpgfhfckdahglylmdga 257
Qy 206 PYF-----SDRYKTPWGLTFNFDGDDQVRKFFLENVEYWFKTFKIDGLRLDAV----- 255
Db 258 ptveyanekdrenyvwg-tanf-dlgkpevrsflisnalfwleyyhdgfrvdavnmly 315
Qy 256 ----HAIFDNS-PKHILQETAEK--AHQLGKFVIAESDLNDPKIVKD-----DCGYKID 302
Db 316 wpndrlyenpyaveflrklneavfaydpnalmiaedstdwkvptaptgyeglgfnykwn 375
Qy 303 AQWVDDF-----HHAVHAFITKEKDYYYQDFGRIEDIEKTFKDVYDGYKYSRYR 352
Db 376 mgwmndmlkymetppyerhrhngvtfsllays-----enfilpshdevvhgkks--- 427
Qy 353 GRTHGAPVGDLPPrKFVFIQNHQVGNRNG---ERLSILTDKTTVYMAATLYILSPYI 409
Db 428 -----linkmpgsyeekfaqlrllygymma-----hpgk 456
Qy 410 PLIFMGEEYETNPFFFSDFDPVLKGVREGRLKENNOMIDPQSEAEFLSKLSW--- 466
Db 457 klifmgnef-----aqfdewkfedeldwvlf 482

Qy 467 --KIDEVLDYKOLINIRKRY 486
Db 483 dfeihrkmdymkeliacykry 504

```

RESULT 14

AAR23787 ID AAR23787 standard; Protein; 658 AA.

XX AAR23787;

XX 04-NOV-1992 (first entry)

XX Heat-resistant pullanase.

XX Enzyme; pTPI.
XX Synthetic.
XX JP04099489-A.
XX 31-MAR-1992.
XX 27-MAR-1990; 90JP-0077230.
XX 27-MAR-1990; 90JP-0077230.
XX (EZAK) EZAKI GLICO CO.
XX (IATR) IATRON LAB INC.
XX WPI; 1992-157263/19.
XX P-PSDB; AAR23787.
XX Heat-resistant pullulanase gene - present recombinant plasmid are new
XX and produced in large amts. using a proper host
XX Claim 3; Page 2-3; 8pp; Japanese.
XX The sequence encoding this protein may be contained in a recombinant
XX plasmid. The heat-resistant pullulanase can be produced in large amts.
XX using a suitable host.
XX Plasmid pTPI contg. the pullulanase gene (Japan Patent Appln. No.
XX 174037/88) was cleaved by various restriction enzymes and inserted
XX into plasmid pV118/119. Single-stranded DNA was prep'd. using helper
XX phage M13K07 and the base sequence determined by the dideoxy method.
XX Sequence 658 AA;

Query Match 12.1%; Score 356; DB 13; Length 658;
Best Local Similarity 22.9%; Pred. No. 4.1e-24;
Matches 141; Conservative 99; Mismatches 200; Indels 176; Gaps 26;

QY 2 FSPGGNI-----EKNKGIFKLWAPVNSVKLKL-----SKKLIPMEKNDGFEVEIDDI 51
DB 44 ffydgdlgwaytrekfkfwaptatevkvklidptekaeajalerlegvwtvtvfgd 103

QY 52 EENLTYSYII---EDKREIPDPASRYQPLGVHDKSLIRTDY-QIIDLQKVKIE----- 101
DB 104 megrytyvvcnavvreaavdp-----yavavsvngeygvviidlakrttekpipp 154

QY 102 -----DLIYELHVGF-----SQEGNFKGVIEK-----LDYKDLGITGI 137
DB 155 itspadaviyemhirdftihpesgvthkgylgltepetkgnptptglsylkelgvtvh 214

QY 138 ELMPPVQFPG-----NRDWCYDGVFLYAVQNTY---GGPW---ELAKLVNEAHR 181
DB 215 ellpfndfagvdrnplqynvgynpihynapegsyatdpndopyarirelkrailhtlqn 274

QY 182 GIAVILDVYVNHIGPEGNYL-----LGLGPYFS-DRYKTPWGLTFNFDRCDCQ--V 230
DB 275 girvmdvynhv-----yirdqssfekivpgyfyfydagpsngtvgndiaserkmv 329

QY 231 RFEILENVEYWFKFIDGLRLDAVIAIFDNSPKHILQETAEKAHOLGFVIAESLNDP 290
DB 330 rkwiidsrvfweeyhvngfrfdlm-giildvetmkavreml-----dtldp 374

QY 291 KIVKDCGVKIDAQWDDFHHAHVAFITKEKDYIYDFGRIEDIEKTFKDFVYDCKYSR 350
DB 375 slivfegwldptp-----isseqkatmqnaeklpri-gyfindyfrdsykgst 421

QY 351 YRGTHGAPVG-----LPPRFVVFIONHD-----OVGN 380
DB 422 fqlpdrfgalgsddrekvkvaisgiknglflhtptqinyveshdhntfwdkneian 481

QY 381 RGNRGERLSILTDKTYLMAATLYLSPIPLIFPMGEYYETNPFPPFFSDPVLIKYR 440
DB 155 GFLYAVQNTYGGPWELAKLVNEAHRKRTAVILDVYVNH---GPEGNYLLGLGPYF--- 208

DB 482 gheseq-----trkrqrklataiavlsgqipflhsqgfyrtk-----qgi- 522
QY 441 EGRLENNQMDPOSEAEFLKSKLSWKIDDEVLD---YKOLINIRKRYNNCKRVEVR 497
DB 523 -----dmsynapds-----inqidwkrksehdvrvyvgllirkrkshraffttea 571
QY 498 EGNCTILIMEKIGI 513
DB 572 nnhlffleptasia 587

RESULT 15
AAR11271
ID AAR11271 standard; Protein; 639 AA.
AC AAR11271;
XX 29-MAY-1991 (first entry)
XX B.stearothermophilus branching enzyme.
XX Branching enzyme; amylaceous substances; transferase;
XX thermostability.
XX Bacillus stearothermophilus.
XX EP418945-A.
XX 27-MAR-1991.
XX 22-AUG-1990; 90EP-0202264.
XX 23-AUG-1989; 89NL-0002128.
XX (COOP-) COOP VERKOOP AVEBE.
XX Kiel JAKW, Boels JM, Beldman G, Venema G;
XX WPI; 1991-088587/13.
XX N-PSDB; AAQ11072.
XX Branching enzyme obt'd. from Bacillus stearothermophilus - has good
XX thermostability, for use in modifying starch-like substances
XX Disclosure; fig 4; 13pp; English.
XX This branching enzyme is derived from B.stearothermophilus and is
XX encoded by the gigB gene. It is stable at high temps. and can be
XX produced in large amts. using recombinant techniques. It is a
XX transferase allowing formation of alpha(1-6) branches in starch-like
XX substances. It is useful for extending the shelf life of such
XX materials and can be used in the prepn. of human and animal foods.
XX Sequence 639 AA;

Query Match 11.6%; Score 342.5; DB 12; Length 639;
Best Local Similarity 24.9%; Pred. No. 7.2e-23;
Matches 135; Conservative 89; Mismatches 197; Indels 121; Gaps 26;

QY 4 FGNIEKNKI-----EKLWAPVNSVKLKL-----KKLIPMEKNDGFEVEIDIEE 53
DB 26 fgahvikkngmvgtrfcvapharevrlvgsfnewngtnfnlmkvsngqvmwifipenle 85

QY 54 NLTYSVII---EKREI---PDPASRYQPLGVHDKSQLIR-TDYOILD-----LGKVKI 100
DB 86 ghlykyeitndngnvlksdpafyafyselrphstasivynikgyqndqtwrrkkqrkryd 145

QY 101 EDLIIVELHVGF---SQEGNF---KGVIEKLDYKDLGITGIELMPVAQFPCNRDMGYD 154
DB 146 qplfiyghfsgwkckedgsfityqemaeeelpyviehgfthlellplvehpfdrsyq 205

QY 155 GFLYAVQNTYGGPWELAKLVNEAHRKRTAVILDVYVNH---GPEGNYLLGLGPYF--- 208

Db 206 gigyysatsrygtbhdmyfidrchqagigvilgwppghfckdshgylmfdgapayeyan 265
Qy 209 -SDRYKTPWGLTFNFDORGCQVRKFILFENVEYWFKFKIDGLRLDAV-----HAI 258
Db 266 mqdrenywg-tanf-dlgkevrsflisnalfmweyfhvdgfrvdavanmlywpnsdvl 323
Qy 259 PDNS-PKHILQEIAEK--AHOLGKFVIAESDLNDPKIVKD--DCG----YKIDAQWVDDF 309
Db 324 ykntyaveflqklnetvfaydnpnlmaedstdwprvtaptgydgglgfnykwnmgwmdi 383
Qy 310 HHAVHAFITTEKDYIYODFGRIEDIEKTFKDVVYDGGKYSRYRGRTHCAPVGDLPPrKFV 369
Db 384 --ltymetppehrkyvhn-----kvtfsllyaysef----- 413
Qy 370 VFIONHQVGNRNGERLSIL----TDKTTYLMAATLYILS-PYIPLFMGEYYETNPF 424
Db 414 ilpfshdevh-gkksllskmpgtyeekfaqlrillygylthpghkklfmggefgq---- 468
Qy 425 FFFSDFSDPVLKGVREGRLKENNQIDPOSEEAFKSLKSWKIDEEVLDYIKOLINIRK 484
Db 469 --fdewkd-----leqldwml-----fdfdmhrnmnyvkvellkcyk 503
Qy 485 RY 486
Db 504 ry 505

Search completed: July 15, 2002, 12:17:34
Job time: 71 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:18:04 : Search time 25.09 Seconds
(without alignments)
541.277 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSGGNIKNKGIFKLWAP.....KLKDELIKVNRGVGVYQLE 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles.pep:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2954	100.0	556	US-08-505-377-1	Sequence 1, Appli
2	2954	100.0	556	US-08-798-269-1	Sequence 1, Appli
3	2954	100.0	556	US-09-055-210-1	Sequence 1, Appli
4	1748.5	59.2	559	US-09-242-690A-15	Sequence 15, Appl
5	969	32.8	589	US-08-399-646-2	Sequence 2, Appli
6	969	32.8	589	US-08-607-321-2	Sequence 2, Appli
7	969	32.8	589	US-08-961-240-2	Sequence 2, Appli
8	969	32.8	589	US-08-605-501-2	Sequence 2, Appli
9	969	32.8	596	US-08-399-646-12	Sequence 12, Appl
10	969	32.8	596	US-08-607-321-12	Sequence 12, Appl
11	969	32.8	596	US-08-961-240-12	Sequence 12, Appl
12	969	32.8	596	US-08-605-501-12	Sequence 12, Appl
13	959	32.5	597	US-08-399-646-4	Sequence 4, Appli
14	959	32.5	597	US-08-607-321-4	Sequence 4, Appli
15	959	32.5	597	US-08-961-240-4	Sequence 4, Appli
16	959	32.5	597	US-08-605-501-4	Sequence 4, Appli
17	959	32.5	598	US-08-399-646-14	Sequence 14, Appl
18	959	32.5	598	US-08-607-321-14	Sequence 14, Appl
19	959	32.5	598	US-08-961-240-14	Sequence 14, Appl
20	959	32.5	598	US-08-605-501-14	Sequence 14, Appl
21	360	12.2	652	US-08-528-026C-4	Sequence 4, Appli
22	319.5	10.8	718	US-09-346-237-10	Sequence 10, Appl
23	310.5	10.5	606	US-09-187-124-2	Sequence 2, Appli
24	300.5	10.2	740	US-08-410-784A-4	Sequence 4, Appli
25	298	10.1	713	US-09-346-237-9	Sequence 9, Appli
26	296.5	10.0	818	US-08-410-784A-2	Sequence 2, Appli
27	296.5	10.0	818	US-09-346-237-11	Sequence 11, Appl

28	292	9.9	829	4	US-09-514-599-6	Sequence 6, Appli
29	292	9.9	862	4	US-09-346-237-1	Sequence 1, Appli
30	284	9.6	893	4	US-09-514-302-4	Sequence 4, Appli
31	284	9.6	1938	4	US-09-514-302-2	Sequence 2, Appli
32	276	9.3	921	4	US-09-514-599-2	Sequence 2, Appli
33	268	9.1	915	4	US-09-346-237-2	Sequence 2, Appli
34	268	9.1	928	4	US-09-514-599-4	Sequence 4, Appli
35	264.5	9.0	820	4	US-09-313-677-21	Sequence 21, Appl
36	264.5	9.0	926	4	US-09-313-677-2	Sequence 2, Appli
37	264.5	9.0	933	4	US-09-313-677-19	Sequence 19, Appl
38	264.5	9.0	967	4	US-09-313-677-17	Sequence 17, Appl
39	250	8.5	928	1	US-08-474-140-11	Sequence 11, Appl
40	250	8.5	928	1	US-08-477-630-11	Sequence 11, Appl
41	250	8.5	928	1	US-08-472-293-11	Sequence 11, Appl
42	250	8.5	928	1	US-08-474-545-11	Sequence 11, Appl
43	250	8.5	928	2	US-08-478-341-11	Sequence 11, Appl
44	250	8.5	928	3	US-08-996-733-11	Sequence 11, Appl
45	249	8.4	588	2	US-08-339-715A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-505-377-1
; Sequence 1, Application US/08505377
; Patent No. 5856146
; GENERAL INFORMATION:
; APPLICANT: MITSUZUMI, Hitoshi
; APPLICANT: KUBOTA, Michio
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME WHICH
; TITLE OF INVENTION: RELEASES TREHALOSE FROM NON-REDUCING SACCHARIDE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,377
; FILING DATE: 21-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 190180/1994
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 109128/1995
; FILING DATE: 21-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP NOT YET RECEIVED
; FILING DATE: 04-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: MITSUZUMI-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; TELEX: 249688
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

Claims copy pep.

US-08-505-377-1

Query Match 100.0%; Score 2954; DB 2; Length 556;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSGGNIERNKGIKFLWAPVNSVKLKKLIPMKNDGFEVEIDDEENLTYSYI 60
Db 1 MFSGGNIERNKGIKFLWAPVNSVKLKKLIPMKNDGFEVEIDDEENLTYSYI 60

Qy 61 IEDKREIPDPASRYQPLGVHDKSLIRTDYQILDGKVKIEDLIYELHVGTFSGQGNFK 120
Db 61 IEDKREIPDPASRYQPLGVHDKSLIRTDYQILDGKVKIEDLIYELHVGTFSGQGNFK 120

Qy 121 GVIEKLDYLDGLGTITELMPVAQFPGNRDVGVDGFLYAVONTYGGPWELAKLVNEAHK 180
Db 121 GVIEKLDYLDGLGTITELMPVAQFPGNRDVGVDGFLYAVONTYGGPWELAKLVNEAHK 180

Qy 181 RGIAVILDVYVNHIGPEGNYLLGLGYPFSDRYKTPWGLTFNFDGDCQVRKFFILENVEY 240
Db 181 RGIAVILDVYVNHIGPEGNYLLGLGYPFSDRYKTPWGLTFNFDGDCQVRKFFILENVEY 240

Qy 241 WFTFKIDGLRLDAVHAIFDNPSPKHILQETAIAKAHQLGKGFVIAESDLNDPKIVKDDCGYK 300
Db 241 WFTFKIDGLRLDAVHAIFDNPSPKHILQETAIAKAHQLGKGFVIAESDLNDPKIVKDDCGYK 300

Qy 301 IDAQWDDFHHAVHAFITKEKDYQYQFGRLEDIEKTFKDVGVYVYDGYKSYRGRTHGAPV 360
Db 301 IDAQWDDFHHAVHAFITKEKDYQYQFGRLEDIEKTFKDVGVYVYDGYKSYRGRTHGAPV 360

Qy 361 IDLPKRVFVFIQNHQDQVNGRNGERLSILTDKTYLMAATLYILSPYILFIENGEEY 420
Db 361 IDLPKRVFVFIQNHQDQVNGRNGERLSILTDKTYLMAATLYILSPYILFIENGEEY 420

Qy 421 TNPPFFSDFSDPVLKGVREGRLKNNQMDPOSEAFKLSKSWKIDEEVDLYYKOLI 480
Db 421 TNPPFFSDFSDPVLKGVREGRLKNNQMDPOSEAFKLSKSWKIDEEVDLYYKOLI 480

Qy 481 NIKRYNNKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLIGIGPPKLLK 540
Db 481 NIKRYNNKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLIGIGPPKLLK 540

RESULT 2
US-08-798-269-1
; Sequence 1, Application us/08798269
; Patent No. 6027918
; GENERAL INFORMATION:
; APPLICANT: MITSUZUMI, Hitoshi
; APPLICANT: KUBOTA, Michio
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME WHICH RELEASES TREHALOSE FROM NON-REDUCING SACCHARIDE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/798,269

Handwritten: **Hand DNA**

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/505,377
FILING DATE: 21-JUL-1995
APPLICATION NUMBER: JP 190180/1994
FILING DATE: 21-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 109128/1995
FILING DATE: 11-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP NOT YET RECEIVED
FILING DATE: 04-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: MITSUZUMI-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
TELEX: 249688
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 556 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-798-269-1

Query Match 100.0%; Score 2954; DB 3; Length 556;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSGGNIERNKGIKFLWAPVNSVKLKKLIPMKNDGFEVEIDDEENLTYSYI 60
Db 1 MFSGGNIERNKGIKFLWAPVNSVKLKKLIPMKNDGFEVEIDDEENLTYSYI 60

Qy 61 IEDKREIPDPASRYQPLGVHDKSLIRTDYQILDGKVKIEDLIYELHVGTFSGQGNFK 120
Db 61 IEDKREIPDPASRYQPLGVHDKSLIRTDYQILDGKVKIEDLIYELHVGTFSGQGNFK 120

Qy 121 GVIEKLDYLDGLGTITELMPVAQFPGNRDVGVDGFLYAVONTYGGPWELAKLVNEAHK 180
Db 121 GVIEKLDYLDGLGTITELMPVAQFPGNRDVGVDGFLYAVONTYGGPWELAKLVNEAHK 180

Qy 181 RGIAVILDVYVNHIGPEGNYLLGLGYPFSDRYKTPWGLTFNFDGDCQVRKFFILENVEY 240
Db 181 RGIAVILDVYVNHIGPEGNYLLGLGYPFSDRYKTPWGLTFNFDGDCQVRKFFILENVEY 240

Qy 241 WFTFKIDGLRLDAVHAIFDNPSPKHILQETAIAKAHQLGKGFVIAESDLNDPKIVKDDCGYK 300
Db 241 WFTFKIDGLRLDAVHAIFDNPSPKHILQETAIAKAHQLGKGFVIAESDLNDPKIVKDDCGYK 300

Qy 301 IDAQWDDFHHAVHAFITKEKDYQYQFGRLEDIEKTFKDVGVYVYDGYKSYRGRTHGAPV 360
Db 301 IDAQWDDFHHAVHAFITKEKDYQYQFGRLEDIEKTFKDVGVYVYDGYKSYRGRTHGAPV 360

Qy 361 IDLPKRVFVFIQNHQDQVNGRNGERLSILTDKTYLMAATLYILSPYILFIENGEEY 420
Db 361 IDLPKRVFVFIQNHQDQVNGRNGERLSILTDKTYLMAATLYILSPYILFIENGEEY 420

Qy 421 TNPPFFSDFSDPVLKGVREGRLKNNQMDPOSEAFKLSKSWKIDEEVDLYYKOLI 480
Db 421 TNPPFFSDFSDPVLKGVREGRLKNNQMDPOSEAFKLSKSWKIDEEVDLYYKOLI 480

Qy 481 NIKRYNNKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLIGIGPPKLLK 540
Db 481 NIKRYNNKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLIGIGPPKLLK 540

Qy 541 DELIKVNRGVGYQLE 556
Db 541 DELIKVNRGVGYQLE 556

RESULT 3
US-09-055-210-1
; Sequence 1, Application US/09055210
; Patent No. 6346394
; GENERAL INFORMATION:
; APPLICANT: MITSUZUMI, Hitoshi
; APPLICANT: KUBOTA, Michio
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME WHICH
; RELEASES TREHALOSE FROM NON-REDUCING SACCHARIDE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; -COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/055,210
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/505,377
; FILING DATE: 21-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 109128/1995
; FILING DATE: 11-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP NOT YET RECEIVED
; FILING DATE: 04-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: MITSUZUMI-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; TELEX: 249688
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-055-210-1

Query Match 100.0%; Score 2954; DB 4; Length 556;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFSFGNIEKNKGIFKLWAPYNSVKLSKLLIPMEKNDGFEVEIDDIENLTSYI 60
DB 1 MFSFGNIEKNKGIFKLWAPYNSVKLSKLLIPMEKNDGFEVEIDDIENLTSYI 60
QY 61 IEDKREIPDASRYQPLGVHDKSOLLRTDYQLDLGKVKIEDLIYELHVGTFSSQGNFK 120
DB 61 IEDKREIPDASRYQPLGVHDKSOLLRTDYQLDLGKVKIEDLIYELHVGTFSSQGNFK 120
QY 121 GVIEKLDYLDLGTITGIELMPVAQPGNDRWDGVDGVFLYAVQNTYGGPWELAKLVNEAHK 180
DB 121 GVIEKLDYLDLGTITGIELMPVAQPGNDRWDGVDGVFLYAVQNTYGGPWELAKLVNEAHK 180
QY 181 RGIAVILDVYVNHIGPEGNYLLGLSPYFSDRYKTPMGLTFNFDRCDDQVRKFILENVEY 240

DB 181 RGIAVILDVYVNHIGPEGNYLLGLSPYFSDRYKTPMGLTFNFDRCDDQVRKFILENVEY 240
QY 241 WFKTEFKIDGLRLDAVHAIFDNPSPKHILQEAIAKAHOLGKFVIAESDLNDPKIVKDDCGYK 300
DB 241 WFKTEFKIDGLRLDAVHAIFDNPSPKHILQEAIAKAHOLGKFVIAESDLNDPKIVKDDCGYK 300
QY 301 IDAQWVDDFHAFITREKDYVYQDFGRIDIEKTFKDFVYDGYKSYRGRTHGAPV 360
DB 301 IDAQWVDDFHAFITREKDYVYQDFGRIDIEKTFKDFVYDGYKSYRGRTHGAPV 360
QY 361 GDLPPRKFVVFIONHDOVNGRNGERLSILTDKTTVYMAATLYLSPYPLIFMGEEY 420
DB 361 GDLPPRKFVVFIONHDOVNGRNGERLSILTDKTTVYMAATLYLSPYPLIFMGEEY 420
QY 421 TNPFFFFSDFSDPVLKIGVREGRLKENNOMIDPQSEAFKLSKLSWKIDDEVLDYKQLI 480
DB 421 TNPFFFFSDFSDPVLKIGVREGRLKENNOMIDPQSEAFKLSKLSWKIDDEVLDYKQLI 480
QY 481 NIKRYNNCKRKVEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540
DB 481 NIKRYNNCKRKVEVRREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540
QY 541 DELIKVNRGVGVYQLE 556
DB 541 DELIKVNRGVGVYQLE 556
RESULT 4
US-09-242-690A-15
; Sequence 15, Application US/09242690A
; Patent No. 6284534
; GENERAL INFORMATION:
; APPLICANT: KONDO, KEIJI
; APPLICANT: MIURA, YUTAKA
; TITLE OF INVENTION: YEAST VECTOR AND METHOD OF PRODUCING PROTEINS USING THE
; FILE REFERENCE: 049441/0118
; CURRENT APPLICATION NUMBER: US/09/242,690A
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/JP97/02924
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: JP 8/241062
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 15
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein
US-09-242-690A-15

Query Match 59.2%; Score 1748.5; DB 4; Length 559;
Best Local Similarity 58.9%; Pred. No. 1.3e-162;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;
QY 3 SFGGNIENKKGIFKLWAPYNSVKLK-LSKKLLIPMEKNDGFEVEIDDIENLTSYII 61
DB 2 TFAKIDGNEVIFTLWAPYQKSVKLVLEKGLYEMERDEKGYTFTINNVKVDRIKYVL 61
QY 62 EDKREIPDASRYQPLGVHDKSOLLRTDYQLDLGKVKIEDLIYELHVGTFSSQGNFK 121
DB 62 DDASEIPDASRYQPLGVHGPSQIIQESKEFNNETLKKEDLIYELHVGTFSTFTFEG 121
QY 122 VIRKLDYLDLGTITGIELMPVAQPGNDRWDGVDGVFLYAVQNTYGGPWELAKLVNEAHK 181
DB 122 VIRKLDYLDLGTITGIELMPVAQPGNDRWDGVDGVFLYAVQNTYGGPWELAKLVNEAHK 181
QY 182 GVIAVILDVYVNHIGPEGNYLLGLSPYFSDRYKTPMGLTFNFDRCDDQVRKFILENVEY 241
DB 182 GLGVILDVYVNHIGPEGNYLLGLSPYFSDRYKTPMGLTFNFDRCDDQVRKFILENVEY 241

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/607,321
;; FILING DATE: 26-FEB-1996
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/399,646
;; FILING DATE: 07-MAR-1995
;; APPLICATION NUMBER: JP 59834
;; FILING DATE: 07-MAR-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 59840
;; FILING DATE: 07-MAR-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BROWDY, Roger L.
;; REGISTRATION NUMBER: 25,618
;; REFERENCE/DOCKET NUMBER: KUBOTA-5
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; TELEX: 248633
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 589 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-607-321-2

Query Match 32.8%; Score 969; DB 1; Length 589;
Best Local Similarity 39.5%; Pred. No. 3.2e-86;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

QY 13 GIFKLPVNSVKLSKLLIPMEKN-----DEGFFEVEIDDEENLTYSYIIE-DKR 65
DB 8 GRFDIWPAGVITLLAGERYEMGRPNPGPADEGWTAADAPTADVDYGYLLDGDDEI 67

QY 66 EIPDPASRYQPLGVHDKSLIRTDYQILDGKVKIEDL-----IYELHVGTFSSQ 115
DB 68 PLPDPTRRQPEGVH---ALSRT---FDPGAHRWDAGWQRELOSGSVIYELHIGTFP 120

QY 116 EGNFGVIEKLDYLDLGTIGTIELMPVAFPGNDRNGYDGVFLYAVONTYGGPWEAKLV 175
DB 121 EGTDLAAAGKLDYLAGLIGDIFELLPVNAFNGTHNNGYDGVQWFAVHEGYGPAAYQRFV 180

QY 176 NEAHKRGIAVILDVYVNHIGPEGNYLLGLGYPFSDRYKTPWGLTFNFDGRCDOVRKFIL 235
DB 181 DAAHAAAGLVIOQVYVNHIGPSGNYLPRYGPYLKHGEGNTWGSVNLDPGSDHVRQYIL 240

QY 236 ENVEYWFKTFKIDGLDVAHAFIFONSFKHILQETAEAKAHQLGK-----FVIAESDLND 289
DB 241 DNVAMWLRDVRVYDGLRDLDAVHALKDERAVHILEEFGALADALSSEGGRLTLIAESDLNN 300

QY 290 PKIV--KDCGKIDQWDDHVAHVAFITKEKDYDYQDFGRIEDIEKTFKDVVYDVK 347
DB 301 PRLLYPRDNGYGLAGWQSDHFAHVHVVNSGETTGYSDFSGLGALAKLVLRDGFPHDGS 360

QY 348 YSRYGRTHGAPV--GDLPRKFVVFQIONHDOVNGRNGERSILTDKTTYLMAATLYIL 405
DB 361 YSFRGRCHRPINFSAVHAPALVVCSSQNHQDIGNRATGDRLSQSLLPYGSLAAVLTIT 420

QY 406 SPYIPLIFMGEYETNPFFFSDFSDPVLKIGVREGRLKENNQ-----IDPQSEEA 458
DB 421 GPFTPMLFNGEYGATTPWQFTSHPEPELGRKATASGRIFERMGWDPAVVPDPDQDPET 480

QY 459 FLKSKLSWK-----IDEEVDLYKQILINIRKYNCKRKYKVEVRENCITLIMEKIGITA 513
DB 481 FTRSKLDWAEASAGDHARLLELYRSLTILR-----RSTPELARLGFADTAV----- 526

QY 514 SFDD 517
DB 527 EFDD 530

RESULT 7
US-08-961-240-2
; Sequence 2, Application US/08961240
; Patent No. 5830715
; GENERAL INFORMATION:
; APPLICANT: KUBOTA, Michio
; APPLICANT: TSUSAKI, Keiji
; APPLICANT: HATTORI, Kazuko
; APPLICANT: SUGIMOTO, Toshiyuki
; TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
; TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,240
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,646
; FILING DATE: 07-MAR-1995
; APPLICATION NUMBER: JP 59834
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 59840
; FILING DATE: 07-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-240-2

Query Match 32.8%; Score 969; DB 2; Length 589;
Best Local Similarity 39.5%; Pred. No. 3.2e-86;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

QY 13 GIFKLPVNSVKLSKLLIPMEKN-----DEGFFEVEIDDEENLTYSYIIE-DKR 65
DB 8 GRFDIWPAGVITLLAGERYEMGRPNPGPADEGWTAADAPTADVDYGYLLDGDDEI 67

QY 66 EIPDPASRYQPLGVHDKSLIRTDYQILDGKVKIEDL-----IYELHVGTFSSQ 115
DB 68 PLPDPTRRQPEGVH---ALSRT---FDPGAHRWDAGWQRELOSGSVIYELHIGTFP 120

QY 116 EGNFGVIEKLDYLDLGTIGTIELMPVAFPGNDRNGYDGVFLYAVONTYGGPWEAKLV 175
DB 121 EGTDLAAAGKLDYLAGLIGDIFELLPVNAFNGTHNNGYDGVQWFAVHEGYGPAAYQRFV 180

QY 176 NEAHKRGIAVILDVYVNHIGPEGNYLLGLGYPFSDRYKTPWGLTFNFDGRCDOVRKFIL 235
DB 181 DAAHAAAGLVIOQVYVNHIGPSGNYLPRYGPYLKHGEGNTWGSVNLDPGSDHVRQYIL 240

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/399,646
FILING DATE: 07-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 596 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
US-08-399-646-12

Query Match 32.8%; Score 969; DB 1; Length 596;
Best Local Similarity 39.5%; Pred. No. 3.2e-86;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

QY 13 GIFKLWAPYVNSVKLSKLLIPMEKN-----DEGFFEVEIDDIENLTYSIIIE-DKR 65
DB 15 GRFDIWAPEAGTVTLGGERYEMGRPGNGPADEGWMTAADAPTADVDYGYLLDGEI 74
QY 66 EIPDPASRYQPLGVHDKSOLIRTDYQILDGLGVKIEDL-----IYELHVGTFSSQ 115
DB 75 PLDPDTRRQPEGVH---ALSRT---FDPGAHRWQDAGWQRELOGSVIYELHIGTFTP 127
QY 116 EGNFKGVIEKLDYKDLGTIGTIELMPVAOPPGNRDNGYDGVFLYAVONTYGGFWELAKLV 175
DB 128 EGTDLAAAGKLDYLAGLIGDIFIELLPVNAFNGTHNGYDGVQWFAVHEGYGPPAAVQRFV 187
QY 176 NEAHKRGIAVILDVYVNHGTPGNYLLGLGYPFSDRYKTPWGLTFNFDRCGDOVRKFL 235
DB 188 DAAHAAGLVQDVVYVNHGTPGNYLLGLGYPFSDRYKTPWGLTFNFDRCGDOVRKFL 235
QY 236 ENVEYWFKFKIDGLRDVAHFIDNSPKHILQEIKAHQLGK-----FVIAESDLND 289
DB 248 DNVAMWLRDVRVDGLRDVAHAKDRAVHILLEFGALADALSSEGGRPPLTLIAESDLNN 307
QY 290 PKIV--KDCGKYKIDAQWDDFHHAVHAFITTEKDYIYODFGRIEDIEKTFKDVFIYDVK 347
DB 308 PRLLYPRDNGYGLAGQWSDDFHHAVHVNVSGETTGYSDFDLSGALAKVLRDGFHDGS 367
QY 348 YSRYRQTHGAPV--GDLPPRFVVIQNHQDQVGNRNGRERLSILTDKTYLMAATLYIL 405
DB 368 YSFRGRCHGRPTNFSAVHPALVWCQNHQDQVGNRNGRERLSILTDKTYLMAATLYIL 427
QY 406 SPYIPLIFMGEYEYENPFFSDFDPVLKGVREGRKNNQM-----IDPQSEEA 458
DB 428 GFTPTMFMGEYEGATTTPQFFTSHEPELGGKATAGRIREFRGMWDPAVVPDQDPET 487
QY 459 FLKSKLSWK-----IDEEVDLYKQLINIKRYNNCKRYKEVRREGNCITLIMEKIGIIA 513
DB 488 FTRSKLDWAEASAGDHARLLELYRSLLTLR-----RSTPELARLGFADTAV----- 533
QY 514 SFDD 517
DB 534 EFDD 537

RESULT 10
US-08-607-321-12
Sequence 12, Application US/08607321
Patent No. 5716813
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Keiji
APPLICANT: HATTORI, Kazuko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/607,321
FILING DATE: 26-FEB-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,646
FILING DATE: 07-MAR-1995
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 596 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-607-321-12

Query Match 32.8%; Score 969; DB 1; Length 596;
Best Local Similarity 39.5%; Pred. No. 3.2e-86;
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;

QY 13 GIFKLWAPYVNSVKLSKLLIPMEKN-----DEGFFEVEIDDIENLTYSIIIE-DKR 65
DB 15 GRFDIWAPEAGTVTLGGERYEMGRPGNGPADEGWMTAADAPTADVDYGYLLDGEI 74
QY 66 EIPDPASRYQPLGVHDKSOLIRTDYQILDGLGVKIEDL-----IYELHVGTFSSQ 115
DB 75 PLDPDTRRQPEGVH---ALSRT---FDPGAHRWQDAGWQRELOGSVIYELHIGTFTP 127
QY 116 EGNFKGVIEKLDYKDLGTIGTIELMPVAOPPGNRDNGYDGVFLYAVONTYGGFWELAKLV 175
DB 128 EGTDLAAAGKLDYLAGLIGDIFIELLPVNAFNGTHNGYDGVQWFAVHEGYGPPAAVQRFV 187
QY 176 NEAHKRGIAVILDVYVNHGTPGNYLLGLGYPFSDRYKTPWGLTFNFDRCGDOVRKFL 235
DB 188 DAAHAAGLVQDVVYVNHGTPGNYLLGLGYPFSDRYKTPWGLTFNFDRCGDOVRKFL 247

RESULT 13

SALES PR
20Z
COMPLAINTS
8T; MISMA
186; INDETS
34; GAPS
10;

QY	15	FKLWAPYNSVKLKUSKLLIPWEK-----NDGFFVEVDDIENLWTSYIIE-DKREI	67
DB	16		68
DB	17		69
DB	18	YDWAPAPAEVTLAGGERYAMQRRRETGPEDAGWTTAAGPTDGNVDYGYLLDGDETPL	77
DB	19		78
QY	68	POPASRYOPLGVHDKASQLI-RTDYQILD--LCKVKLTEDLIYELHVGCTFSOEGNFKGYI	123
DB	69		124
DB	70	POPTRTRQPDGVHIALSRFTDPSAYSQDDAWQGR-ELQGAVIYELHGLTFTPEGTLEAAA	133
QY	124	EKLIDYKLDLGTIGIELMPVAQPGCNROWDGVFLYAVQNTYGGPWEKLAKLVNEAHKRG	181
DB	125		182
DB	137	GKLDYLAGLVDFIELLPVNAFNGTHWNGYDGVQWFAVHEAYGCPYAYORFVDAHAAGL	190
QY	184	AVILDVVVNHIGPEGNYLLGLGPYFSDRYKTPMGLTFFNFDRCDCQVRKFIENYEFWK	241
DB	185		242
DB	197	GVTDVVYVNHLPSCGNLYLPREGPYLKGEGNTWGDVSLDPCGSDHVRVYILDNLAWMLR	251
DB	198		252

; FILING DATE: 07-MAR-1995
; APPLICATION NUMBER: JP 59834
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 59840
; FILING DATE: 07-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KUBOTA-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 597 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-240-4

Query Match 32.5%; Score 959; DB 2; Length 597;
 Best Local Similarity 40.2%; Pred. No. 3.1e-85;
 Matches 202; Conservative 81; Mismatches 186; Indels 34; Gaps 10;

QY	15	FKLWAPVNSVKLKLKIPMEK-----NDGFFVEVEIDDEENLTYSYIIE-DKREI	67
Db	18	YDWWAPNAESVTLLAGERYAMORRAETGPEDAGWTAGAPTDGNVDYGYLLDGDETPL	77
QY	68	PDPSRYQPLGVHDKSOLI--RTDYQILD---LGKVKIEDLIYIELHVGTFSQEGNFKGVI	123
Db	78	PDPRTRQPDGVHLSKRTDPSAYSQDDDAWQGR-ELQGAIVIELHLGTFTEGLEAAA	136
QY	124	EKLIDYLDLIGITGIELMPVAQFPGNDWGDYGVFLYAVQNTYGGPWELAKLYNEAHKGI	183
Db	137	GKLDYLAGLVDFIELLPVNAFNGTINWGYDGVQWFAVHEAYGGPEAYQRFVDAHAAGL	196
QY	184	AVILDVVYVNHIGPEGVNLGLGVFSDRYKTPMGLTFNFDGDCDOVRKFFILENVEYWFK	243
Db	197	GVLDQVYVYHGLSPGNYLPRFGYLYKQGGNTWGDSDVNLDPGSDHVRRYILDNLAMWLR	256
QY	244	TFKIDGLRLDAVHAIFDNPSPKHILQIEIAEKHQLGKEV-----IAESDLNDPKIV--KD	295
Db	257	DVRVDGLRLDAVHALKDERAVHILEDFGALADQISAEVGRPLTLIAESDLNPNRLLYPRD	316
QY	296	DCGYKIDAQWDDFHAVAFITKEKDYIQDEGRIDIEKTFKDVYVDGKYSRYRGRT	355
Db	317	VNSYGLEQWSDDFHVAHVNTGETTYSDYDFSLAALAKVLRDGFHDGYSYSPREH	376
QY	356	HGAPV--GDLPPRKVFVFIQNHQVGNRNGERLSILTDKTLYLMAATLYILSPIPLIF	413
Db	377	HGRPIFSAVHPAALVVCQNHQICGNRATGRLSQTLPYGSIALAAVLTGTFPTMLL	436
QY	414	MGEYETNPFFPFDSPDLIKGVREGRLKENNOM-----IDPQSEEAFLKSKLSW	466
Db	437	MGEYGASTPWQFETSHPEPELGKATAEGRIKEFERMGWDPAVVDPDPQDPETFRSKLDW	496
QY	467	KIDEE-----VLDYYKQLINRK	484
Db	497	AEAAEGDHARLLELYRSLTALRR	519

Search completed: July 15, 2002, 12:18:05
Job time: 97 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:18:43 ; Search time 32.43 Seconds
(without alignments)
1647.415 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSFGNIEKNKGIFKLWAP.....KLKDELKVNRGVGYOLE 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2954	100.0	556	2 JC5132	alpha-amylase (EC
2	1784.5	60.4	561	2 S73087	alpha-amylase (EC
3	1748.5	59.2	558	2 JC5135	alpha-amylase (EC
4	1212.5	41.0	620	2 AH1827	alpha-amylase [imp
5	969	32.8	596	2 JC4697	trehalose trehaloh
6	959	32.5	598	2 S65770	malto-oligosyltreh
7	936	31.7	575	2 JC7727	probable hydrolase
8	900.5	30.5	594	2 AI0673	probable maltocollig
9	866.5	29.3	580	2 G70763	probable glycosyl
10	865.5	29.3	583	2 G83375	probable maltocollig
11	835.5	28.3	600	2 H75516	alpha-amylase Atu5
12	793.5	26.9	595	2 AG3194	probable 1,4-alpha
13	792.5	26.8	601	2 D96001	probable alpha-glu
14	493	16.7	552	2 S19134	hypothetical prote
15	493	16.7	552	2 AH1915	1,4-alpha-glucan b
16	353	11.9	666	2 B56639	glycogen operon pr
17	342.5	11.6	638	2 S18599	alkaline amylopull
18	333.5	11.3	707	2 S77094	alpha-amylase (EC
19	332.5	11.3	1280	2 E95031	pullulanase amyx -
20	331.5	11.2	1256	2 G97902	1,4-alpha-glucan b
21	326	11.0	718	2 G69585	pullulanase amyx -
22	322.5	10.9	422	2 S31839	1,4-alpha-glucan b
23	320	10.8	843	2 H72204	pullulanase - Ther
24	319.5	10.8	718	2 S73088	glycogen operon pr
25	315	10.7	562	2 S03112	alpha-amylase (EC
26	313.5	10.6	627	1 S40048	1,4-alpha-glucan b
27	311.5	10.5	642	2 E98000	1,4-alpha-glucan b
28	310.5	10.5	642	2 G95129	probable glycogen
29	306.5	10.4	666	2 E71565	

30	303.5	10.3	664	2 B86539	glycogen hydrolase
31	303.5	10.3	664	2 B72084	glycosyl hydrolase
32	302	10.2	783	2 B84823	probable isoamylas
33	296.5	10.0	764	2 S47569	1,4-alpha-glucan b
34	296.5	10.0	818	2 T02231	probable isoamylas
35	296	10.0	630	2 D70363	1,4-alpha-glucan b
36	295	10.0	666	2 G81717	glycosyl hydrolase
37	295	10.0	789	2 T01321	probable isoamylas
38	293.5	9.9	717	2 H84057	pullulanase amyx l
39	292.5	9.9	733	2 T49679	probable branching
40	292	9.9	1072	2 A84112	alkaline amylopull
41	289	9.8	720	2 G97229	pullulanase [impor
42	289	9.8	741	1 S34218	1,4-alpha-glucan b
43	286.5	9.7	666	2 D82511	1,4-alpha-glucan b
44	283	9.6	746	2 S74546	glycogen operon pr
45	282	9.5	770	2 S76095	hypothetical prote

ALIGNMENTS

RESULT 1
JC5132

alpha-amylase (EC 3.2.1.1) - Sulfolobus acidocaldarius

N;Alternate names: glycogenase

C:Species: Sulfolobus acidocaldarius

C>Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000

C:Accession: JC5132

R; Kobayashi, K.; Kato, M.; Miura, Y.; Kettoku, M.; Kameda, T.; Iwamatsu, A.

Biosci. Biotechnol. Biochem. 60, 1720-1723, 1996

A; Title: Gene analysis of trehalose-producing enzymes from hyperthermophilic archaea

A; Reference number: JC5131; MUID:97141330

A; Accession: JC5132

A; Status: preliminary; nucleic acid sequence not shown

A; Molecule type: DNA

A; Residues: 1-556 <KOB>

A; Cross-references: DDBJ: D64131; NID: g987704; PIDN: BAA11011.1; PID: g987705

A; Experimental source: strain ATCC33909

C; Comment: This enzyme hydrolyzes glycosyltrehalose to liberate trehalose.

C; Superfamily: trehalose trehalohydrolase

C; Keywords: glycosidase; hydrolase

Query Match 100.0%; Score 2954; DB 2; Length 556;
Best Local Similarity 100.0%; Pred No. 2,3e-194;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSFGNIEKNKGIFKLWAPYVNSVKLSKLIIPMEKNDGEGFEVEIDIEENLTYSYI 60

Db 1 MFSFGNIEKNKGIFKLWAPYVNSVKLSKLIIPMEKNDGEGFEVEIDIEENLTYSYI 60

QY 61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDGKVKIEDLIYELHVGTFSGGNPK 120

Db 61 IEDKREIPDPASRYQPLGVHDKSQLIRTDYQILDGKVKIEDLIYELHVGTFSGGNPK 120

QY 121 GVTEKLDYKLDGITGTELMPPVAQFPGNRDWDGVDGVFLAVQNTYGGPWELAKLVNEAHK 180

Db 121 GVTEKLDYKLDGITGTELMPPVAQFPGNRDWDGVDGVFLAVQNTYGGPWELAKLVNEAHK 180

QY 181 RGIAVILDVYVNHIGPEGNYLLGLGPYFSDRYKTPNGLTFFNFDGRGCDQVRKFIENVEY 240

Db 181 RGIAVILDVYVNHIGPEGNYLLGLGPYFSDRYKTPNGLTFFNFDGRGCDQVRKFIENVEY 240

QY 241 WFKTFKIDGLRLDAVHAIFDNPSPKHILQETAEKAHQLGKFPVIAESDLNDPKIYVDDCGYK 300

Db 241 WFKTFKIDGLRLDAVHAIFDNPSPKHILQETAEKAHQLGKFPVIAESDLNDPKIYVDDCGYK 300

QY 301 IDAQWDDDFHVAHFAITTEKDYVYODFGRIEDIETKFDVFDYDGKYSYRGRTHCAPV 360

Db 301 IDAQWDDDFHVAHFAITTEKDYVYODFGRIEDIETKFDVFDYDGKYSYRGRTHCAPV 360

QY 361 GDLPPRRKVVVFYQIHQDQVGNRNGERLSILTDKTTYIMATLYILSPYPIEFMGEYE 420

|||||

Db 361 GDLPRKVFVFIQNHQVGNRNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEYE 420
Qy 421 TNPEFFSDFSDPVLKGVREGRLKNNQMDPQSEAFKSKLSWKIDEEVLDDYKOLI 480
Db 421 TNPEFFSDFSDPVLKGVREGRLKNNQMDPQSEAFKSKLSWKIDEEVLDDYKOLI 480
Qy 481 NIKRYNNCKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLIGIGPKPKLK 540
Db 481 NIKRYNNCKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLIGIGPKPKLK 540
Qy 541 DELIKVNRGVGVYOLE 556
Db 541 DELIKVNRGVGVYOLE 556
RESULT 2
S73087
alpha-amylase (EC 3.2.1.1) precursor - Sulfolobus solfataricus
N:Alternate names: protein c0620
C:Species: Sulfolobus solfataricus
C:Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
C:Accession: S73087
R:Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A:Title: Organizational characteristics and information content of an archaeal genome: I
A:Reference number: S73076; MUID:97055432
A:Accession: S73087
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-561 <SEN>
A:Cross-references: EMBL:Y08256; NID:gl707679; PIDN:CAA69503.1; PID:gl707699
A:Experimental source: strain p2
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C:Superfamily: trehalose trehalohydrolase
C:Keywords: glycosidase; hydrolase
Query Match 60.4%; Score 1784.5; DB 2; Length 561;
Best Local Similarity 59.2%; Pred. No. 2.4e-114;
Matches 332; Conservative 90; Mismatches 130; Indels 9; Gaps 6;
Qy 3 SFGNIEKNKGIFKLWAPYVNSVKL--LSKKLIPMEKNDEGFVEIDDEENLTYSYII 61
Db 2 TFGYKLEDGVTNFWAPYQKVKLTNLRGIVEMERDDKGYFTITLNNVGRDVKYIL 61
Qy 62 EDKREIPDPASRYQPLGVHDKSOLIRTDYQILDGLGVKIEDLIYELHWGTFQSGNF 119
Db 62 DDNSEVPDPASRYQPEGVHGYSEIISDPDFWDDNSVKVREDLVIELHIGFTSGTF 121
Qy 120 KGVTEKLDYLDLKITGIELMPVAQPPGNRDWGYDGVFLYAVONTYGGPWELAKLVNEAH 179
Db 122 EGVTKLNLKELGVTAIEIMPIAQPPKDWGYDGVFLYAVONSYYGSGPGRKLVNEAH 181
Qy 180 KRGTAVIDVYVNHGPEGNYLLGLGYPFSDRYKTPMGLTFNFDRCDDQVRKEILENVE 239
Db 182 KLGAVILDDVYVNHVGPENYMKLGYPFSEKTKPMGLTFNFDAGSDEVRKEILENVE 241
Qy 240 YWFKTKIDGLRDVAHAIFDNPSPKHILQRIAEKHAQOLGVIAESDLNDPKIV--KDDC 297
Db 242 YWNEFVQDGLRDVAHAIDNPKHILEADIADVHKYKIVIAESDLNDPRVVPNPKKC 301
Qy 298 GYKIDAQWDDFHAVHAFITKEKDYDQFGRIEDIEKTFKDVYVYDGYKSYRGRTHG 357
Db 302 GYNIDAQWDDFHAHAIFLTERGQGYSDFGSIGDIVSKYKDVFIYDGYKSNFRRTKG 361
Qy 358 APVGDLPFRKVFVFIQNHQVGNRNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEE 417
Db 362 KSVGDLGCRFVYVFIQNHQVGNRNGERLIKLVKESYKIAAALYILSPYIPIFMFEE 421
Qy 418 YETNPFFSDFSDPVLKGVREGRLKNNQMDPQSEAFKSKLSWKIDEEVLDDYK 477
Db 422 YGENPYFFSDFSDPKLQGVREGRRRNGEQETDPSDCTFNDKLSWKINDDILSFYK 481

Qy 478 QLINIRKRYN-NCKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLI--GIGF 534
Db 482 SLIRKEYGLACNKLKSVENGNYWLVKNGCLAVVVFVSKSVIEMKYSGLTVLSSNSF 541
Qy 535 PKLKLKDELKLVNRGVGVYOL 555
Db 542 PSQITESK-YELDKGFALYKL 561
RESULT 3
JC5135
alpha-amylase (EC 3.2.1.1) - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
C:Accession: JC5135
R:Kobayashi, K.; Kato, M.; Miura, Y.; Kettoku, M.; Kameda, T.; Iwamatsu, A.
Biosci. Biotechnol. Biochem. 60, 1882-1885, 1996
A:Title: Gene cloning and expression of new trehalose-producing enzymes from the hype
A:Reference number: JC5134; MUID:97141610
A:Accession: JC5135
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <KOB>
A:Cross-references: DDBJ:D641130; NID:9987702; PIDN:BA11010.1; PID:gl838936
A:Experimental source: strain KMI
C:Comment: This enzyme hydrolyzes starch alpha-1,4 linkage adjacent to the alpha-1,1
C:Superfamily: trehalose trehalohydrolase
C:Keywords: glycosidase; hydrolase
Query Match 59.2%; Score 1748.5; DB 2; Length 558;
Best Local Similarity 58.9%; Pred. No. 6.9e-112;
Matches 329; Conservative 87; Mismatches 136; Indels 7; Gaps 5;
Qy 3 SFGNIEKNKGIFKLWAPYVNSVKL--LSKKLIPMEKNDEGFVEIDDEENLTYSYII 61
Db 1 TFAFKIDGNEVIFTWAPYQKVKLVLEKGLYEMERDEKGYFTITLNNVGRDVKYVL 60
Qy 62 EDKREIPDPASRYQPLGVHDKSOLIRTDYQILDGLGVKIEDLIYELHWGTFQSGNF 121
Db 61 DDASEIPDPASRYQPEGVHGYSPQIQESKEFNFTFLKEDLIYELHWGTFPGTFEG 120
Qy 122 VIEKLDYLDLKITGIELMPVAQPPGNRDWGYDGVFLYAVONTYGGPWELAKLVNEAH 181
Db 121 VIEKLDYLDLKITGIELMPVAQPPGNRDWGYDGVFLYAVONSYYGSGPGRKLVDEAHK 180
Qy 182 GIAVILDDVYVNHGPEGNYLLGLGYPFSDRYKTPMGLTFNFDRCDDQVRKEILENVEY 241
Db 181 GLGVILDDVYVNHVGPENYMKLGYPFSEKTKPMGLTFNFDAGSDEVRKEILENVEY 240
Qy 242 FKTEKIDGLRDVAHAIFDNPSPKHILQRIAEKHAQOLGVIAESDLNDPKIV--KDDCY 299
Db 241 IKEYNDGFLRDVAHAIDTSPKHILEADIADVHKYKIVIAESDLNDPRVVPNPKKC 300
Qy 300 KIDAQWDDFHAVHAFITKEKDYDQFGRIEDIEKTFKDVYVYDGYKSYRGRTHG 359
Db 301 NIDAQWDDFHHSHAYLTGERQGYTDFGNLDDIVKSYKDVYDGYKSNFRRTKHGE 360
Qy 360 VGDLPFRKVFVFIQNHQVGNRNGERLSILTDKTTYLMAATLYILSPYIPLIFMGEEY 419
Db 361 VGELDCNCFVYVFIQNHQVGNRNGERLIKLVKESYKIAAALYILSPYIPIFMFEEY 420
Qy 420 ETNPFFSDFSDPVLKGVREGRLKNNQMDPQSEAFKSKLSWKIDEEVLDDYKOL 479
Db 421 EENPFFSDFSDSKLIQGVREGRRKENGQDTPQDESTNASKLSWKIDEEIFSFKIL 480
Qy 480 INIRKRYN-NCKRVKVRREGNCITLIMEKIGIIASFDIVINSKITGNLLIGI--GFGK 536
Db 481 IKRKELSIACDRRVNVNNGENWLIKGREYSLVYFVSKSVIEVYSGTLLSSNSFPQ 540
Qy 537 KKLKDELKLVNRGVGVYOL 555
Db 541 HIEEGK-YEFDKGFALYKL 558

RESULT

AH1827
 alpha-amylase [imported] - Anabaena sp. (strain PCC 7120)
 C:Species: Anabaena sp.
 A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
 C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
 C:Accession: AH1827
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A:Reference number: AB1807; UID:21595285; PMID:11759840
 A:Accession: AH1827
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-620 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BAB77692.1; PID:g17135146; GSPDR:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all0168
 C:Superfamily: trehalose trehalohydrolase

	Query Match	41.0%;	Score 1212.5;	DB 2;	Length 620;
	Best Local Similarity	46.5%;	Pred. No. 3,7e-75;		
	Matches 235;	Conservative	86;	Mismatches	151;
				Indels	33;
	Gaps				7;
QY	15 FKLWAPVNSVK---	LKLSKKLIPEKNDEGFFEVEIDDIENLTYSVIIEDKREIPDPA	71		
Db	1 : : : : : :	: : : : : : : : : : : : : : :			
QY	29 FTWSPTLNSAVAGILKAPEEKLIPLKAQEGYGWKVNDVPGLYRLYQLNDQEAADPA	88			
Db	1 : : : : : :	: : : : : : : : : : : : : : :			
QY	72 SRYOPLGVHDKSQLIRDYQILD--	LGKVKTEDLLIYELHVGTFSOEGNKGVTLEKLDYL	129		
Db	1 : : : : : :	: : : : : : : : : : : : : : :			
QY	89 SQYOPEGVGASQVVVDHKFEWDTKWTSGISLESIMIFELHWGTTTPGETFTTIIPLRDEL	148			
Db	1 : : : : : :	: : : : : : : : : : : : : : :			
QY	130 KDLGITGIEMPVAQFPFN-----	RDMWGCVGLFYAYQNTYGGPWEAKLVNEAHK	180		
Db	: : : : : :	: : : : : : : : : : : : : : :			
QY	149 RELGINAIELMPIAQFPGDDHHIEPDLAYRNWYGDGYPPAVONSYGSGPADLKNEVNACHE	208			
Db	1 : : : : : :	: : : : : : : : : : : : : : :			
QY	181 RGTAVILDVVYNHHTPGPNYLLGLGPVDSBYKYTKPWGLTFNFDDRGCDOVRKFLEINVEY	240			
Db	: : : : : :	: : : : : : : : : : : : : : :			
QY	209 NGTAIVLDVVYNHHTPGPNYMGQAFYTRYTKYPWNAMNFDDDAY SOGVSNRYFIQNALY	268			
Db	: : : : : :	: : : : : : : : : : : : : : :			
QY	241 WFKTFKIDGLRLDAVHAIFONSPKHIILQEIAEKAHQLG-----	KFVIAESDLNDMPKIVK	294		
Db	1 : : : : : :	: : : : : : : : : : : : : : :			
QY	269 WLGEFHIDGLRLDAIQAILYDLGAHFELWELAAEVHNFSOGQTWRKHLLAESDLNNPQIIR	328			
Db	1 : : : : : :	: : : : : : : : : : : : : : :			
QY	295 --DDCGYKIDAQWDDFFHHAVHAFITEKDYYYODFGRIEDIETKFDVFYDGKYSYR	352			
Db	: : : : : :	: : : : : : : : : : : : : : :			
QY	329 PVESGGYGLDAQNSDDPFHHALLHALTDRCQYQDFGKCADLAKAYADTFVYDWRYAPHR	388			
Db	1 : : : : : :	: : : : : : : : : : : : : : :			
QY	353 GRTHGAPVGDLPPIPKFVVFVTONHDQVNRNGERLSILTDKTTYLMATLYILSPYPLI	412			
Db	1 : : : : : :	: : : : : : : : : : : : : : :			
QY	389 KRFGHSICDRPLPSQSVCIONHDQIGNOMGERLSERISFAGUKLAAGAVALLSPYPLLL	448			
Db	1 : : : : : :	: : : : : : : : : : : : : : :			
QY	413 FMGEYYETNPFFFSDFSDPVILKIGVREGRLKE-----	NNQMVIDPOSEAFKSLXSW	466		
Db	: : : : : :	: : : : : : : : : : : : : : :			
QY	449 FMGEYGETAPPYFVSHVSDDLLOIQA VRAGKEEFEEAFHYAEDDPDESATFLRCKLNW	508			
Db	1 : : : : : :	: : : : : : : : : : : : : : :			
QY	467 KIDE-----	VLOYKKOLINIKRY	486		
Db	: : : : : :	: : : : : : : : : : : : : : :			
QY	509 ELRHQGQHVKVLWDWYROLIHRLKTH	533			
Db	1 : : : : : :	: : : : : : : : : : : : : : :			

RESULT

RESULT 3
 JCX697
 trehalose trehalohydrolase (EC 2.4.1.-) - Rhizobium sp.
 C:Species: Rhizobium sp.
 C:date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Jun-2000
 C:Accession: JCX697
 R:Maruta, K.; Hattori, K.; Nakada, T.; Kubota, M.; Sugimoto, T.; Kurimoto, M.

Biosci. Biotechnol. Biochem. 60, 717-720, 1996

A:Title: Cloning and sequencing of trehalose biosynthesis genes from *Rhizobium sp.* M
A:Reference number: JC4696; MUID:96219094

A:Accession: JC4697

A:Molecule type: DNA

A:Residues: 1-596 <MAR>

A:Cross-references: DDBJ:D78001; NID:g1040695; PIDN:BA11187.1; PID:g1040697

C:Comment: This enzyme belongs to the alpha-amylase family. It hydrolyzes the produc-
oligosyl group and trehalose.

C:Genetics:

A:Gene: MTHase

C:Superfamily: trehalose trehalohydrolase

C:Keywords: glycosyltransferase; hexosyltransferase

```
Query Match          32.8%; Score 969; DB 2: Length 596;  
Best Local Similarity 39.5%; Pred. No. 1.6e-58;  
Matches 215; Conservative 74; Mismatches 195; Indels 60; Gaps 12;  
  
QY  13 GIFKWPAPYNSVKLKSLIPMEKN-----DEGFFEVIDDIEENLYSYIIE-DKR 65  
    |||:|||::|||::|:|:||||:||||::|:|:|:|:|:|:  
Db   15 GREFDIWAPEACTVLLAGGERYEMGRPPNGPADGWNWTAADAPTADGVADVGYLLDGDEI 74  
  
QY  66 EIPDPASRYQLPGVHDKSOLIRTDYQILDIGKKYIEDL-----IIYELHVGTFEQ 115  
    |||:|||::|||:|||::|:|:||||:||||::|:|:|:|:|:|:  
Db   75 PLPDPRTFRROPEGVH----ALSRT-----FDFGAHRWDQAGWGRELQGSVYLIELHGTFTP 127  
  
QY  116 EGNFGKVTEIKLDYKLGLGTIGIELMPVAQPFGNRDWCYGDGVELYAVONTYGGPWELAKLV 175  
    |||:|||::|||:|||:|||::|:|:||||:||||::|:|:|:|:|:|:  
Db   128 EGTDLAAAGKLDYLAGLCIDFIELLPVNAENGTHNWGYDGVPFAVHEGYGGPAAYQRV 187  
  
QY  176 NEAHKRGIAVILDVVYNHIGPEGNYLLGLGFYSDRYKTPWGLTFNFDRCDCQVRKFIL 235  
    ||:|:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db   188 DAAHAAGLGVIQDVVYNHLGPSNYLPYRGPYLKHGEGNTWGDVNLDPGSDHVRQYIL 247  
  
QY  236 ENVEYWFKEFKIDGLRDYDAHYAFDNSPKHILQEIAKAHQLOK-----FVTAESDLND 289  
    ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db   248 DNVAWLURDYRVGLRLDAYHALKDERAVHILEFGALADALSSEGGRPILTTLAESDLNN 307  
  
QY  290 PKIV--KDDCGYKIDAQOWDDEFHHAVHAFITTKEDKYYYQDFGRIEDIKTFKDFVFVDGK 347  
    ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db   308 PRLLYPRDVNGYGLAGOWSDFHFHAVHVNVSGETGYISDFSLGALAKVLVRGFFHDGS 367  
  
QY  348 YSRYPGRKTHGAPV--GDLPRPKFVVFTIQNHDOVGNRGNRERLSILTDKTTYLTMAATLYIL 405  
    ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db   368 YSSPRGRCHGRPINFSAVHPAALVVCSONHDQIGNRATGDRLSQSPLPYGSALAAVLTLT 427  
  
QY  406 SPIPILFMGEEYEYNPFNFSSDFSDDLKVGREGLKENQM-----IDPOSEEA 458  
    ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db   428 GPPTFMLFMGEEXGATTPOWFFTSHPEPELGKATAEGRIFERFERMGMWDPAVVPDPQDET 487  
  
QY  459 FLKSKLSWK-----IDEVLDYKCOLINIBKRNCKRKVEVRREGNCITLIMEKIGIIA 513  
    ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db   488 FTRSKLDWAESAAGDHARLLLEIYRSILTLR-----RSTPELARLGFDATAV----- 533  
  
QY  514 SFDD 517  
    |||  
Db   534 EFDD 537  
  
RESULT      6  
S65770  
maltooligosyltrehalose trehalohydrolase - Arthrobacter sp. (strain Q36)  
C:Species: Arthrobacter sp.  
A:Variety: strain Q36  
C:date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000  
C:Accession: S65770  
R:Maruta, K.; Hattori, K.; Nakada, T.; Kubota, M.; Sugimoto, T.; Kurimoto, M.  
Biochim. Biophys. Acta 1289, 10-13, 1996  
A:title: Cloning and sequencing of trehalose biosynthesis genes from Arthrobacter sp  
A:Reference number: S65769; MUID:96195835  
A:Accession: S65770  
A>Status: preliminary
```

A:Molecule type: DNA
A:Residues: 1-598 <MAY>
A:Cross-references: EMBL:D63343; NID:g1255444; PIDN:BAA09668.1; PID:g1255446
C:Superfamily: trehalose trehalohydrolase

Query Match 32.5%; Score 959; DB 2; Length 598;
Best Local Similarity 40.2%; Pred. No. 7.8e-58;
Matches 202; Conservative 81; Mismatches 186; Indels 34; Gaps 10;

Qy 15 FKLWAPYVNSVKLKLIPMEK-----NDEGFEVEIDIEENLTYSIIIE-DKREI 67
Db 19 YDWNAPNASTLLAGGERYAMQRAETGPEADGWTMTAAGAPTDCGNDVDYGLLDGDETP 78
Qy 68 PDPAASRYQPLGVHDKSOLI-RTDQILD--LGKVKIEDLIYELHVGTFSGQEGNFKGVI 123
Db 79 PDPRTRQPCGVHSLRTFDSAYSWQDDAQGR-ELQGVAYIELHGLTFTPEGTLEAA 137
Qy 124 EKLDYLDGLTGITGIELMPVAOPGNRMGVDGVFLYAVQNTYGGFWELAKLVNEAKRGI 183
Db 138 GKLDYLAGLVDFTELLPVNAFNGTHNMGYDGVQVFAVEAYGGPEAYQRFVDAHAAGL 197
Qy 184 AVILDVYVNHGTPGNYLLGPGYFSDRYKTPWGLTFNFDGRCDOVRKFFILENVVWF 243
Db 198 GVIQDVVYVNHGTPGNYLLGPGYFSDRYKTPWGLTFNFDGRCDOVRKFFILENVVWF 257
Qy 244 TFKIDGLRLDAVHAIFDINSRKHILOEIA-----EKAHQLGK--FVIAESDLN 295
Db 258 DYRVDGLRLDAVHAIFDINSRKHILOEIA-----EKAHQLGK--FVIAESDLN 317
Qy 296 DCGYKIDQWDDPHHVAHFATTEKDYKDYQDFGRIEDIEKTFKDYGVYKGRYRGRT 355
Db 318 VNGYGLEQWSDPHHVAHFATTEKDYKDYQDFGRIEDIEKTFKDYGVYKGRYRGRT 377
Qy 356 HGAPV--GDLPRKRVFQIHNHDOVNGRNGERLSILTOKTLYLMAATLYLTPYIPLIF 413
Db 378 HGRFINSAVHPAALVCSQNHQDIGNRATGDRLSQILPYGSLAALAVLTUTGPTFMIL 437
Qy 414 MGEYETNPFFSDFSDPVLKIGVREGRLKENNQ-----IDPQSEAFKSLSW 466
Db 438 MGEYGASTPWQFTTSHPPELGKATAEGRIKEFERMGWDPAVVPDPQDPETFRSLDW 497
Qy 467 KIDEE-----VLDYKQLINIRK 484
Db 498 AEAAGDGHARLLELYRSILTALRR 520

RESULT 7
JC7727
malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) - Arthrobacter ramosus
C:Species: Arthrobacter ramosus
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 31-Dec-2001
C:Accession: JC7727; PC7175
R;Yamamoto, T.; Maruta, K.; Watanabe, H.; Yamashita, H.; Kubota, M.; Fukuda, S.; Kurimoto, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, B.; Yamamoto, T.; Maruta, K.; Watanabe, H.; Yamashita, H.; Kubota, M.; Fukuda, S.; Kurimoto, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr, S.; Moule, S.; O'Gaora, P.
Biocli. Biotechnol. Biochem. 65, 1419-1423, 2001
A:Title: Trehalose-producing operon treZ from Arthrobacter ramosus S34.
A:Reference number: JC7726; MUID:21364352; PMID:11471747
A:Accession: JC7727
A:Molecule type: DNA
A:Residues: 1-575 <YAM>
A:Cross-references: DDBJ:AB045141
A:Experimental source: strain S34
A:Accession: PC7175
A:Molecule type: protein
A:Residues: 1-20;31-50;298-319 <YA2>
C:Comment: This enzyme is involved in producing trehalose from glycogen.
C:Genetics:
A:Gene: treZ
C:Superfamily: trehalose trehalohydrolase
C:keywords: glycosidase; hydrolase

Query Match 31.7%; Score 936; DB 2; Length 575;

Best Local Similarity 38.7%; Pred. No. 2.8e-56;
Matches 198; Conservative 82; Mismatches 177; Indels 54; Gaps 10;

Qy 15 FKLWAPYVNSVKLKLIPMEKNDGFEFE--IDDIEENLTYSIIIEIDKREIPDP 70
Db 5 FPVWAPQAAQVTLVVGGRAELPLTRDENGWALQQPWDGPDLDVYGLVKGKPPADP 64
Qy 71 ASRYQPLGVHDKSOLIIRTDYQILDGKV-----KIEDLIYELHVGTF 114
Db 65 RSLRQPRGVH-----ELGREDFPARYAWGDDGWRGRDLTGAVIYELHVGTF 111
Qy 115 QEGNFKGVIEKLDYLDGLTGITGIELMPVAOPGNRMGVDGVFLYAVQNTYGGFWELAKL 174
Db 112 PEGTLDGSAIRRLDLHLRVGDAVELLPVNAFNGTHNMGYDGVLMVAVHEPYGGPEAYQRF 171
Qy 175 VNEAHKRGIAVLVDVYVNHGTPGNYLLGPGYFSDRYKTPWGLTFNFDGRCDOVRKFI 234
Db 172 VDACHAGLVAQVQVYVNHGTPGNYLLGPGYFSDRYKTPWGLTFNFDGRCDOVRKFI 231
Qy 235 LENVEYFKTFKIDGLRLDAVHAIFDINSRKHILOEIA-----EKAHQLGK--FVIAESDLN 288
Db 232 IDNAVYWLROMHADGLRLDAVHAIFDINSRKHILOEIA-----EKAHQLGK--FVIAESDLN 291
Qy 289 DPKIVKDDC--GYKIDQWDDPHHVAHFATTEKDYKDYQDFGRIEDIEKTFKDYGVYKGRYRGRT 346
Db 292 DPKIIRSAAHGYGLDAQWDDPHHVAHFATTEKDYKDYQDFGRIEDIEKTFKDYGVYKGRYRGRT 351
Qy 347 KYSYRGRTHGAPV--DLPPRKRVFQIHNHDOVNGRNGERLSILTOKTLYLMAATLYL 405
Db 352 TWSSFRHHRGRLDPLPFRLVAFADQHDQVGNRAVGRMSAQVGEGLSALAAALVLL 411
Qy 406 SPYIPLFMGEYETNPFFSDFSDPVLKIGVREGRLKENNQ-----IDPQSEEA 458
Db 412 GPFTPLFMGEYETNPFFSDFSDPVLKIGVREGRLKENNQ-----IDPQSEEA 471
Qy 459 FLKSKLWKIDEE-----VLDYKQLINIRK 484
Db 472 FARSHLWSEPREPHAGLLAFYTDLTALRR 502

RESULT 8
AI0673
Probable hydrolase STY1503 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium, S.; Moule, S.; O'Gaora, P.
C:Species: Salmonella enterica subsp. enterica serovar Typhimurium, S.; Moule, S.; O'Gaora, P.
A:Note: this species has also been called Salmonella typhimurium
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AI0673
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, B.; Yamamoto, T.; Maruta, K.; Watanabe, H.; Yamashita, H.; Kubota, M.; Fukuda, S.; Kurimoto, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr, S.; Moule, S.; O'Gaora, P.
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium, S.; Moule, S.; O'Gaora, P.
A:Reference number: AB0502; PMID:11677608
A:Accession: AI0673
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-594 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01762.1; PID:g16502610; GSPDB:GN001776
C:Genetics:
A:Gene: STY1503
C:Superfamily: trehalose trehalohydrolase

Query Match 30.5%; Score 900.5; DB 2; Length 594;
Best Local Similarity 38.9%; Pred. No. 7.8e-54;
Matches 199; Conservative 79; Mismatches 173; Indels 61; Gaps 13;

Qy 15 FKLWAPYVNSVKLKLIPMEKNDGFEFEIDIEENLTYSIIIEIDKREIPDPASRY 74
Db 22 FRLWATGQKQVMLRLAGKQEMQASGSGWFTLDVSGVTPGTETYNFLVSLDGMVLPDPASRA 81
Qy 75 QPLGVHDKSOLIIRTDYQILDGKV-----KIEDLIYELHVGTFSGQEGNFKGVI 124

Db 82 QKTDVNGPS-----YVDPGSYAWRNTGWKSRWEQAVVYEMHTGTFTPEGTFTATA 134
QY 125 KLDYLKDLGITGIELMPVAFPCNDRWGDYGVFLYAVQNTYGGPWLAKLVNEAHRKRGIA 184
Db 135 KLPYLAELGVTVIEWPVAQFGGERGWGDYGVLLYAPHSAYGTPDPDFKAFIDTAHGYGLS 194
QY 185 VILDVYVNHGPEGNYLLGLSP-YFSDRYKTPWGLTFNFDRCGQOVRKFIENVEYWFK 243
Db 195 VILDVNLHFGPEGNYLLPLAPAFHKKRMTWPWGNGIAYD---VDAVRYIITEAPLYWLT 251
QY 244 TKIDGLRLDAVHAFTDNPSPKILQIEAKAHQGLCKFVIAESDLNDPKI----- 292
Db 252 EYHLGLRLFDAIDQIEDSSARHVLVEIAQRIE-----DITDRPILHLTEDSRNII 302
QY 293 ---VKDDCYK--IDAQWDDHFAHAFITKEKDYDYDFGRIED--IEKTFKDVYVD 345
Db 303 SLHPRDODGNAPLFTAENWDDFNHNVFATGETQYNYNDFADAPKHLARALAEAFYQ 362
QY 346 GKYSRYGRTHGAPVGDLPFRKFVFIQNHQVGNRNGERLSILTD-KTTYLMAATLYI 404
Db 363 GEISPOTGPRGVKSTGQPPVAFVDFIQNHQVGNRAQGRDLITLAGAERTKVLLATL-L 421
QY 405 LSPYIPLIFMGEEYETNFFFSDFSDPVLKGVREGRLKE-----NNQMDPQSEAF 459
Db 422 LSPHIPPLFMGEYGESRPFLEFTDFHGD-LARAVREGRAKEFADHAGENVPDPNAPETF 480
QY 460 LKSKLSWKIDEEV-----LDYKQKILNIRKY 486
Db 481 QRSKLNKQKHSEEGKAWLAFTRELLLLKQKH 512

RESULT 9
G70763
probable maltotoligosyltrehalose trehalohydrolase - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: G70763
R:Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987
A:Accession: G70763
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-380 <COL>
A:Cross-references: GB:274020; GB:AL123456; NID:g3261584; PIDN:CRA98329.1; PID:g1403480
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: glqZ
C:Superfamily: trehalose trehalohydrolase

Query Match 29.3%; Score 866.5; DB 2; Length 580;
Best Local Similarity 37.3%; Pred. No. 1.6e-51;
Matches 187; Conservative 83; Mismatches 194; Indels 37; Gaps 11;
QY 15 FKLWAPYNSVKLSKULIPMEKNDGFEVEIDIEENLYSYIIEKREI-PDPASR 73
Db 4 FRWAPKPAVLVDVNGAVHAMTRSDAGWWHTV-AAPADARYGYLLDDDPVLPDPRSA 62
QY 74 YQPLGVHDKSQ-----LIRTDYQILDGLKVKIEDLIYELHVGTFSQSGNFKGVLE 124
Db 63 ROPDGVHARSQRWEPGPQGAARTD---TGWGRSVSEGAIVYELHIGTFTTAGTFDAALIE 119
QY 125 KLDYLKDLGITGIELMPVAFPCNDRWGDYGVFLYAVQNTYGGPWLAKLVNEAHRKRGIA 184
Db 120 KLDYLDGLIDFVELMPVNSFAGTGWGDYGVLYSVHPEYGGPDGLRFIDACHARRLG 179
QY 185 VILDVYVNHGPEGNYLLGLSPYFSDRYKTPWGLTFNFDRCGQOVRKFIENVEYWFK 244

Db 180 VLIDAVFNHLGFSNGVLPFRGPLYSSA-SNPMGDSGINIAGADSDEVRRHYIDCALRWMD 238
QY 245 FKIDGLRLDAVHAIFDNPSPKILQIEAKAH---OLGK-FVIAESDLNDPKIV--KDD 296
Db 239 FIADGLRLDAVHAIVDVTAVHVLEELANATRWLSQGLRPLSIAETDRNDPLITRPSH 298
QY 297 CGYKIDAQWDDHFAHAFITKEKDYDYDFGRIEDIEKTFKDVYDGKYSRYGRTH 356
Db 299 GGYGITAQWDDTHIHAHTAVSGERQGYADFGSLATLAYTLRNGYFHAGTYSRPRRRH 358
QY 357 GAPV--CDLPFRKFVFIQNHQVGNRNGERLSILTDKTTYLMAATLYLSPIPLIFM 414
Db 359 GRALDTSATPATRLLAYTCTHDPQGNRGLDRPSOYLTTGGQLAKAALTIGSPYTAAMLFM 418
QY 415 GREYETNPFPPFSDFSDPVLKGVREGRLKE-----NNQMDPQSEAFKLSKSNK 467
Db 419 GEWGASSPFQFCSPHELAHSIVAGRKEFAEHWGAADDIPDPQDPQTFQRCCKLNA 478
QY 468 -----IDEEVLDYKOLINIR 483
Db 479 EAGSGEHAHLRHFYRDLIALR 499

RESULT 10
G83375
probable glycosyl hydrolase PA2164 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83375
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337
A:Accession: G83375
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-583 <STO>
A:Cross-references: GB:AE004643; GB:AE004091; NID:g9948178; PIDN:AAG05552.1; GSPDB:G83375
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2164
C:Superfamily: trehalose trehalohydrolase

Query Match 29.3%; Score 865.5; DB 2; Length 583;
Best Local Similarity 38.2%; Pred. No. 1.9e-51;
Matches 199; Conservative 89; Mismatches 166; Indels 67; Gaps 20;
QY 4 FGNIEKN-KGTFKLWAPYNSVKLSK-KLIPMEKNDGFEVEIDIEENLYSYII 61
Db 7 FGAQFGCGNRTCFGLWAPDAREVRVETADGRDWPLSGSGWFEATL-PCPGTRTRYRI 65
QY 62 EDKREIPDPSRYQPLGVHDKSQLIRTDYQILDGLKV--KIE-----DLIYELHVG 111
Db 66 DGRPGVPDPASQFPQDGVHGS-----QVLHDGTYAWKRVDEWRGPRWHEAVIYELHVG 118
QY 112 TFSQENFGKVIKLDYLDGLGITGIELMPVAFPCNDRWGDYGVFLYAVQNTYGGPWL 171
Db 119 LF---GSYAEVERFLPRLVELGVTAVELMPLGEFFGPRRNWGDVLPFPAPASAYGTPQEL 175
QY 172 AKLVNEAHRKGTAVILDVYVNHGPEGNYLLG-LGPFYSDRYKTPWGLTFNFDRCGQO 230
Db 176 KHLIDSAHGCMGLMVFVDYIYNHFGPCNLAQYAAAFRRDRTQPMGQALDF-RRG--EV 232
QY 231 RKFIENVEYWFKTFKIDGLRLDAVHAIFDNPSPKILQIEAKAH-----QLGKFVIAES 285
Db 233 REFFYNALMWLLDYRVDGLRFDVHAIPDSA---FLVEMARLRGAAGPERHVLHVLN 289
QY 286 DUNDPKIVKDDCGYKTDQAOWDDFHAVIAFTKEDKDYIYQDFGR-IEDIEKTFKDVFY 344

Db 290 DNRASLLRQ--GY--DAQWDDGHHALHVLITGENDGYQDYPPRLRLCLARCLAEGFY 345

QY 345 DQKYSRYGRTHGAPVDLPPKRVFFVIONHDQVGNRNGERLSILTDKTTYLMAATLYI 404

Db 346 OGEANRH-GRPRGEPADLAPDAFLFQNHHDQVGNRAFGERLSVLAEPQALRLAIALQL 404

QY 405 LSPYIPLIFMGEEYETNPFFSFSDPVLKIKVREGRLKENNQ-----MID 452

Db 405 LAPMPLIFMGEECAAREPFLLFTDHQGE-LADAVREGRRKEFGFGRFGEGATLASLPD 463

QY 453 POSEAEFLKSK-----LSWKIDDEVLDYKOLINIRKRY 486

Db 464 PNAVETFRSRLACDPAWR-----GFYQLLEIRHEH 498

RESULT 11

H75516

A:Title: Probable maltotigosyltrehalose trehalohydrolase - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: H75516

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: H75516

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-600 <WHI>

A:Cross-references: GB:AE001905; GB:AE000513; NID:g6458144; PIDN:AAF10042.1; PID:g645814

C:Genetics:

A:Map position: 1

C:Superfamily: trehalose trehalohydrolase

Query Match 28.3%; Score 835.5; DB 2; Length 600;

Best Local Similarity 36.8%; Pred. No. 2.2e-49;

Matches 181; Conservative 89; Mismatches 191; Indels 31; Gaps 12;

QY 15 FKLWAPYVNSVKLKLIPMEKNDGFEVEIDDEENLYTYIIEKREIPDPASRY 74

Db 38 FLMTSTARTVAVRVNGTEHVMTSLGGGIYELEL-PVGPARYLFVL-DGVPTDPYARF 95

QY 75 QPLGVHDKSQLI---RTDYQLDLGKVKIEDLIIEYELHVGTFSGEKNFGVIEKLDYLDK 131

Db 96 LPDGVHGEAEVDFGTDFWTDADWHGKGLADCVYEVHVGTFTPPEGYRAAAEKLPYLKE 155

QY 132 LGITGIELMPVAOPGPNRDGMDYGVFLYAVQNTYGGPWELAKLYNEAKHKGIAVILDVY 191

Db 156 LGVTAIQVMPLAAPDQGRGWYDGAFAVPAYGRPEDLMALVDAHRLGLGVFLDVY 215

QY 192 NHIGPEGNYLLGLCP-YFSDRYKTPWGLTFNFDRCQDQVKKFLENVVWFKTFKIDGL 250

Db 216 NHFGPSNLYSSYAPSYFTDRFSAWGMGLDYAE---PHMRRYVTGNARMLRDYHFDGL 272

QY 251 RLDAVHAIFNSPKHILQETAERKAHQLG--KFVIAESDLNDPKIVKDCGKIDAQWDD 308

Db 273 RLDAATPYMTDSETHILTELAAQEHGLGTHLLAEHRNLPDLVTYN---HLDGIWTD 329

QY 309 FHAAVHAFTKEKDYQDF-GRIEDIETKFDKVDYVGVYDGRYSYRGRTH--GAPVGLDP 365

Db 330 FHETRVTLTGEQGYAGYAGGAELAYTIRRGWRYEGQFWAVKGEERHGPSDALEA 389

QY 366 RKFVVFQNHQDVGNRNGERLSILTDKTTYLMAATLYIILSPYIPLIFMGEEYETNPFF 425

Db 390 PNFVFIQNHQDQIGNRPLSERLHQSDGVTLLHEYRGAALLPMTPLLFQGEWAASPTFQ 449

QY 426 FFDSPDSDPVLKIKVREGRLKE-----NNQIMDPQSEAEFLKSLKSWIDE-----EV 472

Db 450 FFSDHAGE-LGOAVSEGRKKKEFGSGFSGEDVDPDQAQOTFLNSKLNWAEREKGGEHART 508

QY 473 LDYYKOLINIRK 484

Db 509 LRLYRDLLRLR 520

RESULT 12

AG3194

A:Title: alpha-amylase Atu5284 [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plasmid

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C:Accession: AG3194

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woc erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; MCCI : Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2333, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AG3194

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-595 <KUR>

A:Cross-references: GB:AE008687; PIDN:AAL45973.1; PID:g17743726; GSPDB:GN00188

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu5284

A:Genome: plasmid

C:Superfamily: trehalose trehalohydrolase

Query Match 26.9%; Score 793.5; DB 2; Length 595;

Best Local Similarity 36.7%; Pred. No. 1.6e-46;

Matches 175; Conservative 76; Mismatches 195; Indels 31; Gaps 11;

QY 15 FKLWAPYVNSVKLKLIPMEKNDGFEVEIDDEENLYTYIIEKREIPDPASRY 74

Db 22 FSIWAPASATVKLNDLNDAEFMDHTAGDGHDTKFPALPGD-RYGFVLADRTFVADPASNR 80

QY 75 QPLGVHDKSQLIRTDYQILD---LGKVKIEDLIIEYELHVGTFSGEKNFGVIEKLDYLDK 131

Db 81 QOEGPRGPSLIIVNHDFAWKPNWKR-PWHEAVYVELHIGTFTPEGTFAAAAEKLEVLAD 139

QY 132 LGITGIELMPVAOPGPNRDGMDYGVFLYAVQNTYGGPWELAKLYNEAKHKGIAVILDVY 191

Db 140 VGITIELMPLATFAGSRGWYGVLFQSPORDYGTPELKAFLDQAHGHGIMVLDVY 199

QY 192 NHIGPEGNYLLGLGYPFSDRYKTPWGLTFNFDRCQDQVKKFLENVVWFKTFKIDGLR 251

Db 200 NHFGPAGNTLQAYAPAFKKHETPWGPAPDENRA--EVRSEFFLQAFYWLTYRFDGLR 256

QY 252 LDV-HAIFNSPKHILQETAERKAHQLG----FVIAESDLN--DPKIVDCCYKIDAQ 304

Db 257 IDAADHLAGDGEVDFTLEMAREVKRTIRNRHVLHVEDARNAASPMTPMADGAILVDAQ 316

QY 305 WVDDEHVAHAFITKEKDYIYQDFR--IEDIEKTFKDVYDGYKYSRYRG-RTHGAPVG 361

Db 317 WNDHFHIVHVAITNEEGGIYEDFASPYENLRSLATGTYVQGEPRPSRNFASGEPFG 376

QY 362 DLPPKRVFVFIQNHQDVGNRNGERLSILTDKTTYLMAATLYIILSPYIPLIFMGEEYET 421

Db 377 HLPPHREVFNLHNDQAGNRLGERLALPPPLFGTLEALLCQTPPLVFMGDEHGA 436

QY 422 NPFFSFSDPVLKIKVREGRLKE-----NNQIMDPQSEAEFLKSLW 466

Db 437 NPFFSFSDHPHNRQEIR-NRLKAESFQGLPPPDASQVMVMDPDQHTMQLTSLKW 492

RESULT 13

D96001

Probable 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) [imported] - Sinorhizobium m

C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
C:Accession: D96001
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: D96001
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <CUR>
A:Cross-references: GB:AL591985; PIDN:CAC49676.1; PID:gl5141163; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: glgB2; SMD21447
A:Genome: plasmid
C:Superfamily: trehalose trehalohydrolase
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 26.8%; Score 792.5; DB 2; Length 601;
Best Local Similarity 36.9%; Pred. NO. 1.9e-46;
Matches 192; Conservative 79; Mismatches 199; Indels 51; Gaps 15;
QY 3 SGGN-IEKNKGIFKLWAPVNSVKLKSLLIPMEKNDGFEVEIDIEENLTYSYII 61
DB 13 SWGANFIDSTCRFLWAPDERVDVLGGVHKQSLDGGWFEITL-AAKAGERYCFRL 71
QY 62 EKREITPDPSRYQLGVHDKSLI---RTDYLQILDGLVKIEDLIYELHVGTFSEQN 118
DB 72 ADGTEVADPASSAQEREASSTIVVDHAAVENQTSWRGRPWEAEVISELHVGCTPECT 131
QY 119 FKGVEIKLDYLDLGTITELMPVAQPCGNRDWYDGVFLYAVQNTYGGPWELAKLVNEA 178
DB 132 FRAAIELRPLAGAGITALEIMPVAQPGVGRGWYDGVFLYAPHNAYGKPDOLKALVDAA 191
QY 179 HKRGTAIVLDVYVNHGPGNVLGLGP-YFSDRYKTPWGLTFNFDRCDDQVRKFLFN 237
DB 192 HSLGLTVLDVYVNHGPGNVLGLGP-YFSDRYKTPWGLTFNFDRCDDQVRKFLFN 248
QY 238 VEYWFRTFKIDGLRDVAHAFDNPSPKHL---QEIABK-AHQLGKGFVIAESDLMDPKI 292
DB 249 ALYWLGHFRFDGLRDATQIRDTTKPHELVALEHEVREAFERQIHLVLEDAHRRSLL 308
QY 293 VKDDCYKI--DAQWDDPHHVAHFATKEKYVYODF-----GRTEIEKTFKDVVYD 345
DB 309 ORDASGARMFLDAAWMDLHNLHVATGETKGHYLFADPEWK---IRSALEGFVAP 365
QY 346 GYSRYRGTHGAPYGD---LPRKFVVFIONHDQVNGRNGERLSILTDKTYLMAATL 402
DB 366 AKEDNF-----SPEGSRARVPPQGRVNFQNHDDQIGNRFAGERLASLQEDSLRLAAM 419
QY 403 YILSPYIPLIFMGEYEYENPFFFSDFDPVLKIGVREGRL-----KENNM 450
DB 420 HMLTPQIPLLFMGEYGETQPYFFSDYOGET-AAAIRLGRDEAENFGCLPEGKTMDL 478
QY 451 IDPQSEAFKSKLSWK-----IDEVLDTYKQLINIRKY 486
DB 479 PDPPLDDPVAGSKLRNRRATSPAGERHLAYVRDLAVIRQRH 519
RESULT 14
S19134
probable alpha-glucanotransferase - Anabaena variabilis

C:Species: Anabaena variabilis
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
C:Accession: S19134
R:Sato, N.
Plant Mol. Biol. 18, 165-170, 1992
A:Title: Cloning of a low-temperature-induced gene lti2 from the cyanobacterium Anaba
A:Reference number: S19133; MUID:92119230
A:Accession: S19134
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-552 <SAT>
A:Cross-references: EMBL:X59279; NID:g39252; PIDN:CAA41970.1; PID:g39253
C:Superfamily: neopullulanase; alpha-amylase core homology

Query Match 16.7%; Score 493; DB 2; Length 552;
Best Local Similarity 28.6%; Pred. NO. 5.3e-26;
Matches 159; Conservative 76; Mismatches 142; Indels 178; Gaps 26;
QY 15 FKLWAPYVNSVKLKS---KKLIPMEKNDGFE--VEIDIEENL-----TYSYIIEDK 64
DB 7 FTLFAPYKGAALIASFSWQEIIPMKKGDDGYFRTTVELEDCTGYQYKFRVQTRSFEEED 66
QY 65 R-EIPDPASR--YQPLGVHDKSQLIRTDYQILDGLGKVKIED-----LIIEHLVGT 112
DB 67 QWVDVTPYATDIDESSGKDNSTARIKDGEKIVDTVYVWQHDDKPLPADHELVIYELHVG 126
QY 113 FS-----QEGNFKGVIEKLDYLDLGTITELMPVAQPCGNRDWYDGVFLYAVQNTY 165
DB 127 FSGGDDPYARGKYKHVIEKLDYLCELGINALELPPVKPEYDGYSGWYNPRYFFATESY 186
QY 166 GGPWELAKLVNEAHKRGIAVILDVYVNH---SDLNDRKIVKDCGKIDAOVDDFHHA 312
DB 187 GSTADLKLVDCEHQGRGIRLIMDGIYNHSEASSPLTQIDHDYWHHEPRDPNN----- 240
QY 206 PYFSORYKTPWGLTNFD--DRGCD--QVRKFEILENVEYWFRTFKIDGLRDVAHAFDNP 261
DB 241 -----WGPEFNHYEDENLETYPARKFIGDTVRYVWGEYHLDGIRGYDAAHQIANY 290
QY 262 SPKH-ILQEIASKAQHQLGKFVIAE-----SDLNDRKIVKDCGKIDAOVDDFHHA 312
DB 291 DFWMHIAQAKTAKTAGAKPNVNAEHIPETTSITNLGCP-----MDGCWHDSFYHT 340
QY 313 VHAFITKEDYIYQDFGRIDIEKTFKDVYDGYKSYRGRTHGAPVGLDPPRFVFEI 372
DB 341 IRAHIC-----GDTFDE-ILKQVI--DPKROGELGATN-----VVNVL 376
QY 373 QNHD-----QVNGRNGERLSILTDK-----TTLMAATLYILSPYIPLIFMGEY 418
DB 377 TNHDDHIMVELGNR-----EIFHDEAFPRAKLGTAILMTAV-----GVPLIMWGEF 424
QY 419 YETNPFFFSDFSDPVLKIGVREGRLKENNMIDPQSEAFKLSKLSW-----KIDEEVL 473
DB 425 GEYKP-----KQDDQ-----SKIDTWLLGNDLRSFL 451
QY 474 DYKQOLINIRKRYNN 488
DB 452 DYHKGILGLRK--NN 464

RESULT 15
AH1915
hypothetical protein all0875 [imported] - Anabaena sp. (strain PCC 7120)
C:Species: Anabaena sp.
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AH1915
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriga
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata
DNA Res. 8, 205-213, 2001
A:Title: Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH1915

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-552 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA072832.1; PID:g17130220; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0875
C:Superfamily: neopullulanase; alpha-amylase core homology

Query Match 16.78; Score 493; DB 2; Length 552;
Best Local Similarity 28.68; Pred. No. 5.3e-26;
Matches 159; Conservative 76; Mismatches 142; Indels 178; Gaps 26;

QY 15 FKLWAPVNSVKLKL---KKLIPMEKNDGFFE--VEIDDIEENL--/-TYSYIIEDK 64
DB 7 FTLEAPYNKGAALIASDSQWQEPKMGKDDGYFRTTVELEDGYQYKFRVQTRSWFFED 66
QY 65 R--EIPDASR--YQPLGVHDKSQLIRTYQILDGLGVKIED-----LIYELHVGT 112
DB 67 QWVDVTPYATDIDESSGKDSIARIKDGKIVDTYVWQHDDKPLPADHELVIYELHVGD 126
QY 113 FS-----QEGNFKGVIEKLDYKLDGIGTIELMPVAQPGNPDWGYDGVFLYAVONTY 165
DB 127 FSGGEDDPYARGKYKHVIEKLDYLCELGINAIELLPVKEYPGDYSWGNPNRYFFATESY 186
QY 166 GGPWELAKLVNEAKRGIAVILDVYVYNI-----GPEGNYLLGLG 205
DB 187 GSTADLKLKLVDECHQIRIIMDGIYNHSEASSPLTQIDHDYVYHHEPRDPDN----- 240
QY 206 PYFSDRYKTPWGLTFNFD--DRGCD--QVRKFILENVEYWFKFKIDGLRLDAVHAIFDN 261
DB 241 -----WGPEFNYEHVDENLETYPARKFIGDTVRYWVGEVHLDGIRYDAARQIANY 290
QY 262 SPKH-ILQETAERKAHQLGKFVIAB-----SDLNDPKIVKDDCGYKIDAQWVDDPHHA 312
DB 291 DFHWIAQEAQKTAGAKPFYNVAEHIPEITTSITNLDGP-----MDGCWHDSEYHT 340
QY 313 VHAFITREKDYVYQDFGRIEDIEKTFKDFVYDGKYSRYGRTHGAPVGDLPPrKFVVF 372
DB 341 IKAHIC-----GDTFDE-CLKDVI--DPKRGQFLGATN-----VVNYL 376
QY 373 QNHD-----QVGNRNGERLSILTDK-----TTLMAATLYILSPYIPLIFMGEY 418
DB 377 TNDHHDHIMVELGNR-----EIFHDEAFRAKLGTAIIMTAV-----GVPLIWMGEF 424
QY 419 YETNPFFFSDFDPVLKGVREGRLENQMDPQSEEAFLKSLW-----KIDEEVL 473
DB 425 GEYKP-----KQDQ-----SKIDWTLGNDLNRSLF 451
QY 474 DYKQOLINIRKRYNN 488
DB 452 DYHKGGLIGLRK--NN 464

Search completed: July 15, 2002, 12:18:44
Job time: 136 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:24:32 ; Search time 19.24 Seconds
(without alignments)
1118.923 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSPGGNIKNKGIFKLWAP.....KLKKDELIKVNRGVGYQLQ 556

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	Length	ID	Description
1	969	32.8	596	1	TREZ_RHISP	Q53238 rhizobium s
2	959	32.5	598	1	TREZ_ARTSP	Q44316 arthrobacte
3	936	31.7	575	1	TREZ_ARTRM	Q9aj66 arthrobacte
4	890.5	30.1	580	1	TREZ_BREHE	O52520 brevbacter
5	866.5	29.3	580	1	TREZ_MVCTU	Q10769 mycobacteri
6	353	11.9	666	1	GLGB_BACCL	P30537 bacillus ca
7	342.5	11.6	639	1	GLGB_BACST	P30538 bacillus st
8	320	10.8	843	1	PUBA_THEMEA	O33840 thermotoga
9	315	10.7	562	1	AMY2_DICTH	P14898 dictyoglomu
10	313.5	10.6	627	1	GLGB_BACSU	P39118 bacillus su
11	296.5	10.0	764	1	GLGB_STRAU	P52980 streptomyce
12	282	9.5	770	1	GLGB_SVNY3	P52981 synecocyst
13	276.5	9.4	659	1	GLGX_HAEIN	P45178 haemophilus
14	273	9.2	731	1	GLGB_MVCTU	Q10625 mycobacteri
15	270	9.1	702	1	GLGB_HUMAN	Q04446 homo sapien
16	269	9.1	734	1	GLGB_AGRTU	P52979 agrobacteri
17	267	9.0	639	1	GLGB_BUTFI	P30539 butyrivibri
18	266	9.0	730	1	GLGB_HAEIN	P45177 haemophilus
19	265	9.0	520	1	AMY_BACME	P20845 bacillus me
20	264.5	9.0	773	1	GLGB_SYN7	P16954 synecococc
21	262.5	8.9	498	1	AMY3_DICTH	P14899 dictyoglomu
22	261.5	8.9	558	1	O16G_BACCE	P21332 bacillus ce
23	257	8.7	561	1	TREC_BACSU	P39795 bacillus su
24	254.5	8.6	562	1	O16G_BACTR	P29094 bacillus th
25	252.5	8.5	512	1	AMY1_DEBOC	P19269 debaryomyce
26	250.5	8.5	1196	1	AMYB_FAEPO	P21543 paenibacill
27	249	8.4	588	1	NEPU_BACST	P38940 bacillus st
28	243.5	8.2	478	1	YDD1_SCHPO	Q10427 schizosacch
29	243	8.2	561	1	O16G_BACSU	O06994 bacillus su
30	239.5	8.1	574	1	CDAS_THEET	P29964 thermoanaer
31	237	8.0	494	1	ADAT_SACFI	P21567 saccharomyc
32	234	7.9	555	1	O16G_BACCO	Q45101 bacillus co
33	234	7.9	1475	1	APU_THEYT	P16950 t amyllopull

34	231	7.8	1481	1	APU_THEET	P38939 t amyllopull
35	229.5	7.8	508	1	O16G_BACSP	P29093 bacillus sp
36	228	7.7	799	1	GLGB_MAIZE	Q08047 zea mays (m
37	225.5	7.6	498	1	AMYA_ASPAW	Q02905 aspergillus
38	225.5	7.6	499	1	AMYA_ASPOR	P10529 aspergillus
39	225.5	7.6	499	1	AMYB_ASPAW	Q02906 aspergillus
40	225.5	7.6	569	1	MALT_CANAL	Q02751 candida alb
41	224.5	7.6	535	1	DEXB_STRPN	Q54796 streptococc
42	224	7.6	586	1	AMYM_BACAD	P32818 bacillus ac
43	224	7.6	728	1	GLGB_ECOLI	P07762 escherichia
44	223.5	7.6	499	1	AMY_ASPSH	P30292 aspergillus
45	223.5	7.6	777	1	ISOA_FLASP	O32611 flavobacter

ALIGNMENTS

RESULT	1
TREZ_RHISP	
ID	TREZ_RHISP
AC	Q53238
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE	alpha-D-[(1->4)-alpha-D-glucan]trehalose trehalohydrolase)
DE	(Maltooligosyl trehalose trehalohydrolase).
GN	TREZ.
OS	Rhizobium sp. (strain M-11).
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC	Rhizobiaceae; Rhizobium.
OX	NCBI_TaxID=391;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=96219094; PubMed=8829547;
RA	Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
RT	"Cloning and sequencing of trehalose biosynthesis genes from Rhizobium
RT	sp. M-11."
RL	Biosci. Biotechnol. Biochem. 60:717-720(1996).
CC	- - CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC	linkage in 4-alpha-D-[(1->4)-alpha-D-glucan]trehalose to
CC	yield trehalose and alpha-(1->4)-D-glucan.
CC	- - PATHWAY: Trehalose biosynthesis.
CC	- - SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC	KNOWN AS THE ALPHA-AMYLASE FAMILY
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL; D78001; BAA11187.1; -
DR	InterPro; IPR000461; Alpha amylase.
DR	Pfam; PF00128; alpha-amylase; 1.
KW	Hydrolase; Glycosidase.
FT	ACT_SITE 265 265
FT	BY SIMILARITY.
SQ	SEQUENCE 596 AA; 65262 MW; 73E8A0AE0534DDCD CRC64;

Query Match	32.8%;	Score	969;	DB 1;	Length	596;
Best Local Similarity	39.5%;	Pred.	No. 1e-60;	Indels	60;	Gaps 12;
Matches	215;	Conservative	74;	Mismatches	195;	
QY	13	GFIFKLWAPVSVSKSLKSLIPMEKN-----DEGFFEVEIDDIENITYSYIIF-DKR	65			
Db	15	GRFDIWAPEGTVTLLAGGERVEMGRPGNGPADEGWTAADAPTADVDYGYLLDGEI	74			
QY	66	ETPDASRYOPLGVHDKSOLIFDYOILDGLGVKVIEDL-----TIYELHVGTFSQ	115			
Db	75	PLPDPRTRRQPEGVH---ALSRT-----FDPGAHRWQDAGWQRELQGSVIYELHIGTFP	127			

```
QY 116 EGNEFKGVIEKLDYKDLGIGTIELMPVAQFPGNRDWDGVDGFLYAVQNTYGGPWELAKLV 175
D 116 EGTDAAGAKLDYLAGLIDFIELLPVNAFNGTHWGYDGVQWFAVHEGYGPAAYQRFV 187
QY 176 NEAHKRGTAVIDVYNNHIGPEGNVLLGLGYPFSDRYKTPWGLTFNFDRCDOVRKFL 235
D 176 DAAHAAGLVYQDVYNNHIGPSGNVLPYGYLKHGEGNTWGDVSNLDGPGSDHVRQYIL 247
QY 236 ENVEYWFRTFKIDGLRLDAVHAIFDNPSPKHILQETAEKAKHQLGR-----FVIAESDLND 289
D 236 DNVAMLRDVRDGLRLDAVHALKDERAVHILEEFGALADALSSEGGRPLTIAESDLNN 307
QY 290 PKIV--KDDCYKIDAQVDDFHAVHAFITKEKDYDYQDQGRIEDIEKTFKQVVFVYDG 347
D 290 PRLLYPRDNGYGLAGQWSDDFHVAHVNVSGETTYGYSDFDSLGAALAKVLKRDGFFHDGS 367
QY 348 YSRVGRTHGAPV--GDLPPRKFFVFTQNHDOVNGRNGERLSILTDKTYLMAATLYIL 405
D 348 YSSFRGRCHGRPINFSVHPAALVVCSONHDOIGNRATGDRLSQSLPYGSLALAAVUTLT 427
QY 406 SPYIPLFMGEYETNPFFFSDFSDPVLTKGVREGRLKNNOM-----IDPQSEEA 458
D 406 GPFTPLFMGEYCATTPWQFFTSHPPELKGATAGRIREFERGWDPAVVPDPQDPET 487
QY 459 FLKSLKSWK-----IDEVLDYKOLINIRKRYNCKRVKVRREGNCITLIMEKIGIIA 513
D 459 FTRSKLDWAEASAGDHARLLELYRSLTLR-----RSTPELRLGADTAV----- 533
QY 514 SFDD 517
D 514 EFDD 537

RESULT 2
TREZ_ARTSP
ID TREZ_ARTSP STANDARD; PRT; 598 AA.
AC Q44316;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE alpha-D-[(1->4)-alpha-D-glucano]trehalose trehalohydrolase)
DE (Malto)oligosyl trehalose trehalohydrolase).
GN TREZ.
OS Arthrobacter sp. (strain Q36).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcales; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=1667;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96195835; PubMed=8605217;
RA Maruta K., Hattori K., Nakada T., Kubota M., Sugimoto T., Kurimoto M.;
RT "Cloning and sequencing of trehalose biosynthesis genes from
RT Arthrobacter sp. Q36.";
RL Biochim. Biophys. Acta 1289:10-13(1996).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC linkage in 4-alpha-D-[(1->4)-alpha-D-glucano]trehalose to
CC yield trehalose and alpha-(1->4)-D-glucan.
CC -1- PATHWAY: Trehalose biosynthesis.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: D63343; BAA09668.1; -
DR InterPro; IPR00461; Alpha_amylase.
```

```
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Hydrolase; Glycosidase.
FT ACT_SITE 267 BY SIMILARITY.
SQ SEQUENCE 598 AA; 65831 MW; 8B5C610AD3766947 CRC64;
```

Query Match 32.5%; Score 959; DB 1; Length 598;
Best Local Similarity 40.2%; Pred. No. 5.1e-60;
Matches 202; Conservative 81; Mismatches 186; Indels 34; Gaps 10;

```
QY 15 FKLHAPVYNSVKLSKLLIPMEK-----NDEGFEVEIDDEENITYSVIIE-DKREI 67
D 15 YDWWAPNAESVTLGAGERYAMQRAETGPDAGWTAAGAPTNGNDYDGLDGETPL 78
QY 68 PDASRYQPLGVHDKSOLI--RTDQIILD---LGKVKIEDLIIIVELHVGTSQEGNFKGVI 123
D 68 PDPRTRQPDGVHLSRTFDFPSYQWDDAWQGR-ELQGAIVYIELHGTFTPEGTLEAAA 137
QY 124 EKLDYLDKDLGITGTELMFVAQFPGNRDWDGVDGFLYAVQNTYGGPWELAKLVNEAHKRG 183
D 124 GKLQYLAGLVDFIELLPVNAFNGTHWGYDGVQWFAVHEAGGPEAYQRFVDAHAAGL 197
QY 184 AVLDVYNNHIGPEGNVLLGLGYPFSDRYKTPWGLTFNFDRCDOVRKFLILENVEYWF 243
D 184 GVTDVYNNHIGPSGNVLPFRGYPYKQEGNTWGDVSNLDGPGSDHVRRLDNLAMWLR 257
QY 244 TFKIDGLRLDAVHAIFDNPSPKHILQETAEKAKHQLGVF-----IAESDLNDPKIV--KD 295
D 244 DYRVDGLRLDAVHALKDERAVHILEEFGALADQISAEVGRPLTIAESDLNNPKLLYPRD 317
QY 296 DCGYKIDAQVDDFHAVHAFITKEKDYDYQDQGRIEDIEKTFKQVVFVYDGKYSRYGRGT 355
D 296 VNGYGLGQWSDDFHVAHVNVGTGETTYGYSDFDSLAAALAKVLKRDGFFHDSYSSFRERH 377
QY 356 HGAPV--GDLPPRKFFVFTQNHDOVNGRNGERLSILTDKTYLMAATLYILSPYIPLIF 413
D 356 HGRPINFSVHPAALVVCSONHDOIGNRATGDRLSQSLPYGSLALAAVLTGPTPMLL 437
QY 414 MGEYETNPFFFSDFSDPVLTKGVREGRLKNNOM-----IDPQSEAFKLSLW 466
D 414 MGEYGAFTPWQFFTSHPPELKGATAGRIKEFERMGWDPAVVPDPQDPFTRSKLDW 497
QY 467 KIDEE-----VLDYKOLINIRK 484
D 467 AEAEGDHARLLELYRSLTLR 520

RESULT 3
TREZ_ARTRM
ID TREZ_ARTRM STANDARD; PRT; 575 AA.
AC Q9AJN6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141) (MTHase) (4-
DE alpha-D-[(1->4)-alpha-D-glucano]trehalose trehalohydrolase)
DE (Malto)oligosyl trehalose trehalohydrolase).
GN TREZ.
OS Arthrobacter ramosus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcales; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=1672;
RN [1]
RP SEQUENCE FROM N.A.
RX Yamanoto T., Maruta K., Watanabe H., Yamashita H., Kubota M.,
RA Fukuda S., Kurimoto M.;
RT "Trehalose producing operon trez from Arthrobacter ramosus S34.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(1->4)-D-glucosidic
CC linkage in 4-alpha-D-[(1->4)-alpha-D-glucano]trehalose to
```



```
Matches 187; Conservative 83; Mismatches 194; Indels 37; Gaps 11;

QY 15 FKLWAPYVNSVKLSKLLIPMEKNDGFEVEIDDEENLTYSVITIEDKREI--PDPASR 73
   |::||| |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
Db 4 FRVWAPRALVRLDNGAVHAMTSAAGWHTTV--AAPADARYGLLDDDDTVLPDRSA 62
   |::||| |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::

QY 74 YQPLGVHDKSQ-----LIRTDYQILDGLKVKIEDLIYELHVGTFSGQKFGVIE 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 RQPDGVHARSWRPEPGFGAARDT---TGWPGRSVEGAVIYELHIGTFTAGTDAIE 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 125 KLDYLKDLGITGIELMPVAPPGPNRDWGDVFLYAVQNTYGGPWLAKLVNEAHKRGIA 184
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 KLDYLDLGDIDFVELMPVNSFAGTRGWDVGLVYSVHEPYGGPDGLVRFIDACHARRLG 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 185 VILDVYNHIGPEGNYLLGLPGYFSDRYKTPWGLTFNFDGDCDQVRKFFILENVEYWFKT 244
   |:: |:: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 VLIDAVFNHLGSPSNYLPFRGPLYSSA--SNPWGDCINAGADSDEVRYIIDCALRMWRD 238
   |:: |:: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 245 FKIDGLRLDAVHAIFDNSPKHILQEIKAH-----QLGK--FVIAESDLNDPKTV--KDD 296
   | ||||| ||||| |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
Db 239 FHADGLRLDAVHALVDTTAVHVLLELANATRWLSGQLGRPLSLIAETDRNDPRLITRPSH 298
   | ||||| ||||| |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::

QY 297 CGYKIDAGWDDDEHHAHAFITREKDYVYODFGRIEDIEKTFKDFVYVDGKYSYRGRTV 356
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 299 GGYGITAQWDDDIHHAHTAVSGERGQYADFGLSLATLAYTLRNGYFHAGTYSFRFRRH 358
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 357 GAPV--GDLPPKPVFVFIQNHDPDVGNGRNGERLSILTDKTTYLMAATLYILSPYIPIEM 414
   |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
Db 359 GRALDTSATPRLTAYTCTHDQVGNRAGLGRPSQYLTGGQLAKAALTGSPYATMLFM 418
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 415 GEEYETNPFPPFSDPVLIRKVGREGRLKE-----NNQIDPOSEAFKLSKLSWK 467
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 419 GEWGASSPFQFCSHPELAHSTVAGRKEEFAEHWAAADDPDPQDPQTFQRCKLNWA 478
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 468 -----IDEEVLVDYKQLINIR 483
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 479 EAGSGEHARLHFRYDLIALR 499
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
GLGB_BACCL STANDARD; PRT; 566 AA.
AC P30537;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
DE enzyme).
DE GN GLGB.
OS Bacillus caldolyticus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93208370; PubMed=1296817;
RA Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;
RT "The glgB gene from the thermophile Bacillus caldolyticus encodes a
RT thermolabile branching enzyme.";
RL DNA Seq. 3:221-232(1992).
CC -|- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC glycogen.
CC -|- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC -|- MISCELLANEOUS: OPTIMAL ACTIVITY AT APPROXIMATELY 39 DEGREES
CC CELSIUS.
CC -|- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL; 274020; CAA98329.1;
CC EMBL; AE007027; AAK45880.1; ALT_INIT.
CC TIGR; MT1613;
CC Tubercullist; Rv1562c;
CC InterPro; IPR000461; Alpha_amylase.
CC InterPro; IPR004193; isoamylase.N.
CC Pfam; PF00128; alpha-amylase; 1.
CC Pfam; PF02922; isoamylase.N; 1.
CC Hydrolase; Glycosidase; Complete proteome.
FT ACT_SITE 247 247 BY SIMILARITY.
SQ SEQUENCE 580 AA; 64076 MW; 506190468F44B862 CRC64;

Query Match 29.3%; Score 866.5; DB 1; Length 580;
Best Local Similarity 37.3%; Pred. No. 1.5e-53;
```



```
Db 414 ILPSHDEVH-GKSLLSKMPGTYEKFQALRLLYGYLLTHPGKLLFLMGGERGQ----- 468
Qy 425 FFFSDFSPVLKGVREGRLKENNOMIDPQSEAEFLSKLSWKIDEEVDLYKKOLINRK 484
Db 469 -FDEWKD-----LQOLDWML-----FDFDMHRNNMNVKELLCYK 503
Qy 485 RY 486
Db 504 RY 505

RESULT 8
PULA_THEME
ID PULA_THEME STANDARD; PRT; 843 AA.
AC O33840;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pullulanase precursor (EC 3.2.1.41) (Alpha-dextrin endo-1,6-alpha-
GN glucosidase) (Pullulan 6-glucanohydrolase).
OS PULA OR TM1845.
OC Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=98115241; PubMed=9453151;
RA Bibel M., Brett C., Gosslar U., Kriegshauser G., Liebl W.;
RT "Isolation and analysis of genes for amylolytic enzymes of the
RL hyperthermophilic bacterium Thermotoga maritima.";
RN FEMS Microbiol. Lett. 158:9-15(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Hart D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RL genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of (1->6)-alpha-D-glucosidic
CC linkages in pullulan and in amylopectin and glycogen, and the
CC alpha- and beta-limit dextrins of amylopectin and glycogen.
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ001087; CAA04522.1; -
DR EMBL; AE001821; AAD36907.1; -
DR TIGR; TM1845; -
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; Isoamylase.N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase.N; 1.
KW Hydrolase; Glycosidase; Signal; Complete proteome.
FT SIGNAL 1
FT CHAIN 19 POTENTIAL.
FT ACT_SITE 20 843 PULLULANASE.
FT ACT_SITE 535 535 BY SIMILARITY.
FT ACT_SITE 564 564 BY SIMILARITY.
FT ACT_SITE 652 652 BY SIMILARITY.
SQ SEQUENCE 843 AA; 96261 MW; C42DDE233D54FE77 CRC64;
```

```
Query Match 10.8%; Score 320; DB 1; Length 843;
Best Local Similarity 23.4%; Pred. No. 6.3e-15;
Matches 143; Conservative 88; Mismatches 185; Indels 194; Gaps 31;

Qy 5 GGNIEKNKGIFKLMAPYVNSVKLKS-----KLIPMEKNDGEFFVEIDDIENLT 56
Db 226 GAVYSPEKTIFRVWSPYSKWVLLFKNGEDTEPYQVYVNMVEYKGVWEAVGDLGVF 285
Qy 57 YSIIEDK---REIPDPASRYQPLGVHDKSQ-----LIRTDQ--ILDLG-KVK-TEDL 103
Db 286 YLQLENYGKIRTVDPYSK-----AVYANSKKSAAVWLARTNPGCWENDRGPKIEGEDA 341
Qy 104 IYELHVGTF-----QEGNFK---GVIEKLDYKLDLIGITIELMPVAQ 144
Db 342 IYELHIADITGLNSGVKNKGLYLGLTEENTKPGGVTTGLSHLVGLVTHVHILFFD 401
Qy 145 F-PGNR-----DWGYDGVFLYAV-----ONTYGGPWELAKLYNEAHKRGIAV 185
Db 402 FYTGDELDKDFEKYYNNGYD-PYLFMVPEGRYSTDPKNPHTRIREVKEVMVKALHKHIGV 460
Qy 186 ILDVVYVNHIGPEGNYLLGLG-----PYESDRY-KTPWGLTFNFDRCDOV----- 230
Db 461 IMDVFPFH-----TYGIGELSAFDQTVPIYFYRIDKT--GAYLN--ESGCGNVIASER 509
Qy 231 ---RKFILENVYWFKTFKIDGLRLDAVHAIFDNPSKHILQIEAIAEAKHQLGKFVIAESDL 287
Db 510 PMRKRFIVDTVYVWKEYHYDGFREFDOMGLI---DKTMLEVERALHKI----- 555
Qy 288 NDPKIV-----KDCGYKIDAQWVDFHHAHV--AFITKEKDYVYODFG 329
Db 556 -DPTIILYGEPPWGWGAPIRFGKSDVAGTHVAAPFNDEFDAIRGVSFNPVSKGVFMGGY 614
Qy 330 RIEDIEKTFKDFVYDGKYSRYGRTHGAPVGDLPFRKVFVEIQNHQDVGNRNGERLSI 389
Db 615 KETKIRKGVGSINIDGKLKISFA-----LDPETINYAACHDN-----HT 655
Qy 390 LTKDTTYL-----MAATLYILSPYIPLIFMGGEYETNPFFPFS 429
Db 656 LWDK-NYLAADKADKKKEWTEELKNAQKLAGAILTSQGVFPFLHGGQDFCTKN-FDNS 713
Qy 430 FSDPVLTKGVREGRLKENNOMIDPQSEAEFLSKLSWKIDEEVDLYKKOLINRK----- 484
Db 714 YNAPISNG-----FDYERKLQFI-----DVFENYHKGLIKLRKEHPAF 751
Qy 485 RYNNCKRVKE 494
Db 752 RLKNAEIEKK 761

RESULT 9
AMY2_DICTH
ID AMY2_DICTH STANDARD; PRT; 562 AA.
AC P14898;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Alpha-amylase 2 (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).
GN AMYB.
OS Dictyoglomus thermophilum.
OC Bacteria; Dictyoglomus group; Dictyoglomus.
OX NCBI_TaxID=14;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-8.
RC STRAIN=H-6-12;
RX MEDLINE=88329076; PubMed=2458257;
RA Horinouchi S., Fukusumi S., Ohshima T., Beppu T.;
RT "Cloning and expression in Escherichia coli of two additional amylase
RT genes of a strictly anaerobic thermophile, Dictyoglomus thermophilum,
RT and their nucleotide sequences with extremely low
RT guanine-plus-cytosine contents.";
RL Eur. J. Biochem. 176:243-253(1988).
```



```
Db 26 FGSVRELNGSGYEFVWAPASVRVAGDPSNGSGEHHVHRVNDNGINTFLFPGIGE 85
Qy 54 NLTSY-IIEKREI---PDPAASYOPLGVH-----DKSQLRTIDYQ 91
Db 86 KERYKYEIVTNGEIRLAKDPYAIYSEVRPNTASLTLDLEGYSWQDKWQKQAKTIYE 145
Qy 92 ILDLGVKIEDLIYELHVGTFSGRG-----NFKGVIEKL-DYLDKDLGITGIELMPVAQF 145
Db 146 -----KPVFIYELHLSGSKKHSRGRHSYKELSTQILPIYIKKHGFTHEILLPVVEH 196
Qy 146 PGNRDWGVGYFLYAVONTYGGPWEAKLVNEAKHKGITAVILDVYVYNI--GPESNYLLG 203
Db 197 PYDRSWGTOGYSTPSRFGPPHDLMKFVDECHOQNGVILWVPGHCFKDAHGLYMF 256
Qy 204 LGPYES-----DRYKTPWGLTFNFDRCDOVRKFIENVEYWFKEIDGLRLDAVHAI 258
Db 257 GEPLYEYKEERDRENNLWG-TANF-DLGKPEVHSEFLISNLYWAEFYHIDGFRVDVANI 314
Qy 259 F---DNSPKH---ILQETAERKAHQGLK-----FVIAESDLNDPKI---VKDDCG-----YK 300
Db 315 LYWPNQDERHNPYAVDPLKLNQTMREAYPHVMMIAEDSTEWQVGTGAEEGGLGPHYK 374
Qy 301 IDAOWVDDFHAVHAFITKEDYQ--DFGRIEDIEKTFKDVFYD-----CKYSRYRGR 354
Db 375 WNMGMNDVLKYMET-PPERRHCHQLISFSLLYAFSEHFLPFSDHDEVYVYKKS 429
Qy 355 THGAPVGLDPKRFKVFVFTQNHQDVGNGRNGERLSILTDTKTYLMAATLYILSPYIPLIFM 414
Db 430 -----LNKMPGDYQKQFAQ-----YRLLGLVMTVHPGKLLIFM 462
Qy 415 GEYYETNPFFSDFSDPVLKIGVREGRLKENNQID--POSBEA-----FLKS 462
Db 463 GSEFAQ-----FDEWKD-----TEQLDWFDFSFPMHOKASVFTQDLRLRFYOKS 505
Qy 463 KLSWKIDE-----EVLDDY--KOLINIRKRYNNCKRYKEVREGNCITLINEKIGITASP 515
Db 506 KILYEHDIRAASFEDWVHNDQSFISFIRY-----QKHG-----EALVLIICNF 550
Qy 516 DDIVINSKITG 526
Db 551 TPVVVHYQVDVG 561

RESULT 11
GLGB_STRAU STANDARD; PRT; 764 AA.
AC P52980;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE 1.4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
enzyme).
GN GLGB.
OS Streptomyces aureofaciens.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1894;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM 3239 / ATCC 10762;
RX MEDLINE=94347823; PubMed=8068720;
RA Honerova D., Kormanec J.;
RT "Cloning of the putative glycogen branching enzyme gene, glgB, from
Streptomyces aureofaciens."
RL Biochim. Biophys. Acta 1200:334-336(1994).
CC -1- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
glycogen.
CC -1- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
KNOWN AS THE ALPHA-AMYLASE FAMILY.
-----
CC
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L11647; AAA67437.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR004193; Isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase.
FT ACT_SITE 440 440 BY SIMILARITY.
FT ACT_SITE 493 493 BY SIMILARITY.
FT ACT_SITE 561 561 BY SIMILARITY.
SQ SEQUENCE 764 AA; 85325 MW; 6B45482B4A268ACF CRC64;

Query Match 10.0%; Score 296.5; DB 1; Length 764;
Best Local Similarity 24.6%; Pred. No. 2:5e-13;
Matches 125; Conservative 88; Mismatches 177; Indels 119; Gaps 27;

Qy 3 SFGGNIENKGI-----FKLWAPYVNSVKLSKKL-----IPMEK-NDEGFFEVEIDD 50
Db 162 ALGSOPMEHQCVAGTRFTVWAP--NALGVRVTGDFSYWDVAVAYPMRSLGASGVWELFLPG 219
Qy 51 IEENLTSYII---EDKREI-PDPASRYOPLGVHDKSQLRTIDYQIIDL-----GKVKI 100
Db 220 VAEGALYKYEITRDGGRTLRADPMARYAEVPPANASIVTASRYEWQDAENMARRALAP 279
Qy 101 ED--LIIVELHVGTFSGEGNFKGVIEKLD-YLKDLGITGIELMPVAOPPGNRDMGYDGVF 157
Db 280 HQAPMSVVELHLSWRPGLSYRLAEOLPAYKELGTHVELMPVAEHPGGSGYQVTG 339
Qy 158 LYAVONTYGGPWEAKLVNEAKHKGIAVILDVYVYNHIGPEGNYLLGLGPYFSDR-----YK 213
Db 340 FYAPTRMGTPDDPRFLVDALHRAIGIVIVDWVPAHF-PRDDWALA---EFDGRPLYEHQ 395
Qy 214 TP-----WGLTFNFDRCDOVRKFIENVEYWFKEIDGLRLDAVHAFI----- 259
Db 396 DPRRAAHPDWG-TLEF-DYGRKEVRNFLVANAVYWCQEEFHVGLDRAVASMLYLDYSRD 453
Qy 260 --DNSPK-----HILQETAERKAHQ--GKFFVIAESDLNDPKIVK--DDCGYKID 302
Db 454 EGDWSPNAHGGREDLDAVALLQENNAVYRFPQVVTIAEESTAWDGVTRPTDSGGLGFG 513
Qy 303 AQWVDDFHAVHAFITKE---KDYVQDFGRIEDIEKTFKDVFYVDGKYSRYGRTHGAP 359
Db 514 LKWNMGWMDTLRYVSKPEPVHRKYHHDM-----TFGMVYAFS----- 551
Qy 360 VGDLPPRRKVVVFQNHQDVGNGRNGERLSILT-----DKTYLMAATLYILSPYIPL 411
Db 552 -----ENFVLPI-SHDEI---VHGKR-SLVSKMGPDMWQOORATHRAYLGFMAHPKQL 600
Qy 412 IFMGEEYYETN-----PFFFSDFSDP 433
Db 601 LFMGQEAQGSWSETYGPDWVWLDSSYP 629

RESULT 12
GLGB_SYNY3
ID GLGB_SYNY3 STANDARD; PRT; 770 AA.
AC P52981;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 1.4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
enzyme).
GN GLGB OR SLL0158.
OS Synechocystis sp. (strain PCC 6803).
```

[illegible]

Db 375 HGYOVGNPPSYFAEMNDREDDLCRFLWKSGETGAFAPAGSSDLFKKNDRLPHTTL 434
Qy 337 -----TFKDVYVYDGKYSRYGRTHGAPVGDLPKRFVFTQNDHQVGNRNGERL 387
Db 435 NFITAHDGFTLKDLVSYNQKHNETNGENRGRNE-----NYSYNH---GVEGSTESL 484
Qy 388 -----SIITDKTYYLMAATL--YILSPYIPLIFMGEYYETNPFFFSDFSDPVLKQVR 440
Db 485 SEPOKSAVENNRTPAQSGLLMSLLLANGTPLLADGDFGNTQ----- 526
Qy 441 EGRLEKNQMDPOSEEAFLKLSW-KIDDEVLDYYKQLNIRKRYNCKR 491
Db 527 ----YGNNAAYQDNEITWLK-----WANFNEELFELTKTIALRKQIGSLNK 570

RESULT 14
GLGB_MYCTU STANDARD; PRT; 731 AA.
AC Q10625;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen
branching enzyme).
GN GLGB OR RV1326C OR MT1368 OR MTCY130.11C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S. T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaita F.,
Badoock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies K., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
Hornsbly T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Stulson J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.B., Haft D., Hickey E.,
Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
glycogen.
CC -1- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z73902; CAA98090.1; -;
DR EMBL; AE007010; AAK45632.1; -;
DR TIGR; MT1368; -;

DR TubercuList; RV1326C; -;
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW Complete proteome.
FT ACT_SITE 411 411 BY SIMILARITY.
FT ACT_SITE 464 464 BY SIMILARITY.
FT ACT_SITE 532 532 BY SIMILARITY.
FT CONFLICT 214 214 P -> A (IN REF. 2).
FT CONFLICT 223 223 P -> Q (IN REF. 2).
SQ SEQUENCE 731 AA; 81729 MW; EE2BFEF765352617 CRC64;

Query Match 9.2%; Score 273; DB 1; Length 731;
Best Local Similarity 24.7%; Pred. No. le-11;
Matches 118; Conservative 71; Mismatches 168; Indels 120; Gaps 23;

Qy 15 FKLWAPYVNSVKL-----KLSKKLIPME-KNDEGFFVEIDDIENLTYSY-----II 61
Db 149 FAVWAPNAKGVSLIGEFNGWNGHEAPRVLPSPGWELFWDFDPCDGLYKFRVHGADV 208
Qy 62 EDKREIPDPASRYQPLGVHDKSQLIRTDYQILD-----LGKVKIEDLIYELHVGTF 113
Db 209 TDR---ADPPAFGTEVPPTASRVTSDDYTWGDDDDMMAGRALRNPVNEAMSTYEVHLGSW 265
Qy 114 SOEGNFKGVIEKL-DYLDKDLGITIELMPVAQFPNGNRDWDGYDVLAYQNTYGGPWELA 172
Db 266 RPLSYRQLARELTDYIVDOGFTHVELLPVAEHPFAGSWGVOYTSYAPTFRFGTDPDDR 325
Qy 173 KLVNEAHKRGITAVILDVYVNHIGPEGNYLLGLGVFSDRYK-TP-----WG 217
Db 326 ALVDALHQAGICVIVDWPAPH-PKDAWALG-----RFDCTPLYEHSDDPKRGQLDWG 377
Qy 218 LTFNFDGRGCDQVRKFLFENVEYWFKFKIDGLRLDA-----VHAI 258
Db 378 -TYVF-DGGRPEVRNFLVANALYWLQEFHIDGLKRVDAVMSLYLDYSRPEGGWTPNVHGG 435
Qy 259 FDN-SPRHLQIEIAEKHQL--GKVFIAE-----SDLNDPKIVKDDCYKIDAQWVDDFH 310
Db 436 RENLEAVQFLOEMNATAHKVAPGIVTAEESTPWSGVTRPTNIG---GLGFSMKNNMGWM 492
Qy 311 HAVHAFITKE---KDYVYQDFGRIEDIEKTFKDFVVDGKYSRYGRTHGAPVGDLPKPR 367
Db 493 HTLDYVSRDPVYRSYHHH-----EMTFSLYASENY----- 525
Qy 368 FVFTIQNHQVGNRNGERLSILTDKTTYLMAATLYILSPY-----IPLIFMGEY 418
Db 526 --VLPLSHDEYVH-GKGTLMGRMPG--NNHVKAAGLRSLAYAWAHPGKQLLFMGQEF 578

RESULT 15
GLGB_HUMAN STANDARD; PRT; 702 AA.
AC Q04446;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching
enzyme) (Brancher enzyme).
GN GBE1.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93216700; PubMed=8463281;
RA Thon V.J., Khalil M., Cannon J.F.;
RT "Isolation of human glycogen branching enzyme cDNAs by screening
complementation in yeast.";

J. Biol. Chem. 268:7509-7513(1993).
-!- FUNCTION: REQUIRED FOR SUFFICIENT GLYCOGEN ACCUMULATION. THE
CC ALPHA 1-6 BRANCHES OF GLYCOGEN PLAY AN IMPORTANT ROLE IN
CC INCREASING THE SOLUBILITY OF THE MOLECULE AND, CONSEQUENTLY, IN
CC REDUCING THE OSMOTIC PRESSURE WITHIN CELLS.
CC -!- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
CC glycogen.
CC -!- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER.
CC -!- TISSUE SPECIFICITY: HIGHEST LEVELS FOUND IN LIVER AND MUSCLE.
CC -!- DISEASE: DEFECTS IN GBE1 ARE THE CAUSE OF GLYCOGEN STORAGE DISEASE
CC IV (GSD-IV) (ALSO KNOWN AS ANDERSEN'S DISEASE); A RARE FORM OF
CC GLYCOGENOSIS CHARACTERIZED BY THE ACCUMULATION OF ABNORMALLY
CC STRUCTURED GLYCOGEN THAT RESULTS IN EARLY ONSET HEPATIC CIRRHOSIS,
CC CARDIAC ARREST AND NEUROMUSCULAR DISEASES. MOST CHILDREN WITH THIS
CC CONDITION DIE BEFORE TWO YEARS OF AGE. NO TREATMENT APART FROM
CC LIVER TRANSPLANTATION HAS BEEN FOUND TO PREVENT PROGRESSION OF THE
CC DISEASE.
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L07956; AAA58642.1; -
DR MIM; 232500; -
DR InterPro; IPR000461; Alpha-amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N.1.
KW Glycogen biosynthesis; transferase; Glycosyltransferase;
KW Glycogen storage disease.
FT ACT_SITE 357 357 BY SIMILARITY.
FT ACT_SITE 412 412 BY SIMILARITY.
FT ACT_SITE 481 481 BY SIMILARITY.
SQ SEQUENCE 702 AA; 80443 MW; 4DF3AA8D365A8FE3 CRC64;

Query Match 9.1%; Score 270; DB 1; Length 702;
Best Local Similarity 21.9%; Pred. No. 1.6e-11;
Matches 147; Conservative 90; Mismatches 202; Indels 232; Gaps 29;

QY 3 SFGGNTKNGKIF-KUWAPYVNSVKL-----LSKKLIPMEKNDEGFEVEIDIE----- 52
DB 75 SFGVHRCADGGLYSKEWAPGAEVLTGDFNGWNPFSPYKLDYKWKELYIPPKQNSV 134
QY 53 -----ENLTYSYILEDKREIPDPASRYOPLG 78
DB 135 LVPHGSKLVVTSKSGEILYRISPAKYVVRGDNVNDWIHW-----PEHSYEFK--- 187
QY 79 VHDKSQILRTDQILDGKVKIEDLIYELHVGTFSGQGNFKG-----VTEKLDYLDLGI 134
DB 188 -HSRPK-----KPSRLIYESHVGISSHEGKVASYKFTCNVLPRIKGLGY 232
QY 135 TGIELMPVAQFPGNRDWGVDGVFLYAVQNTYGGPWEAKLVNEAHKRGIAVILDVY----- 190
DB 233 NCITQMAIMEHAYASFGYQITSFFAASSRYGTPEELQELVDVTAHSMGIVLLDVVHSHA 292
QY 191 -----YNHIGPEGNYLILGLPYFSDRYKTPWGLTFNFDRCDCQVRKF 233
DB 293 SKNSADGLNMFDTDCYFHSGRGTHDLWDSRLF-----YSSW-----EVLRF 337
QY 234 ILENVEYFKTKIDGLRLDAVHAIFDNSPKHILQIAEKAHOLGK-FVIAESDLNDPKI 292
DB 338 LLSNIKWLEERFDGFRDGVTSMLYHH-----HGVGGFSGDYSEYFGLQV 385
QY 293 VKDCGCKIDAQWDDFHHAHAF-----ITKEKD-----YYYQDFGRI 331

DB 386 DEDATYLMLA-----NHLVHTLCPDSTITAEDVSGMPALCSPISQGGGGFYRLAMAI 439
QY 332 ED-----IEKFKDV-----FVYDGVKSVRYGRTHGAPVGDLPKRKFVVFVFIQNHDOVGNR 381
DB 440 PDKWIQLLKEFKDEDNMGDIVYTLNRRYL-----EKCIAYAESHDQ----- 482
QY 382 GNGERLSILTDKTT--YLMAATLY-----ILSPYIPLI-----FMGE EY 419
DB 483 -----ALVGDKSLAFWLMDAEMYTNMSVLTPTPVIDRGIQLHKMIRLITHGLGEGYL 536
QY 420 ETNPFFFSDFSDPVLKGVREGRLKENNMIDPQSEEAFLKSKLSWKIDEEVLDYKOL 479
DB 537 N-----FMGNEFGHPWLDPPKRG-----NN-----ESYHAYARQFHLTDDDLLRYKFL 580
QY 480 INIRKRYNNCKR-----VKEVRREGNCITIMEKIGIIASFDDIVINSKITGNL 528
DB 581 NNFDRDMNRLEERYGWLAAPOAYVSE-KHEGNKI-TAFERAGLLFIEN--FHPSKSYTDY 636
QY 529 LIGIGFPKKLK 539
DB 637 RVGTALPGKFK 647

Search completed: July 15, 2002, 12:24:34
Job time: 386 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:24:08 ; Search time 55.94 Seconds
(without alignments)
1719.435 Million cell updates/sec

Title: US-09-298-924-8
Perfect score: 2954
Sequence: 1 MFSPGGNIKNKGIFKLWAP.....KLKRDLEIKVNRGVGVYQLE 556

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_19:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2954	100.0	556	1 Q53641	Q53641 sulfolobus
2	1784.5	60.4	561	17 P95867	P95867 sulfolobus
3	1748.5	59.2	559	1 Q55088	Q55088 sulfolobus
4	1711.5	57.9	559	1 Q9UWN9	Q9UWN9 sulfolobus
5	1290	43.7	359	17 Q973H4	Q973H4 sulfolobus
6	942.5	31.9	581	2 Q9ADI5	Q9ADI5 streptomyces
7	865.5	29.3	583	16 Q911V1	Q911V1 pseudomonas
8	835.5	28.3	600	16 Q9RX51	Q9RX51 deinococcus
9	792.5	26.8	601	16 Q92U63	Q92U63 rhizobium m
10	645	21.8	217	17 Q973H5	Q973H5 sulfolobus
11	493	16.7	552	2 Q44528	Q44528 anabaena va
12	420	14.2	440	2 Q93Q35	Q93Q35 myxococcus
13	360	12.2	652	2 Q59242	Q59242 bacillus st
14	333.5	11.3	707	16 P73608	P73608 synechocyst
15	332.5	11.3	1280	16 Q97SQ7	Q97SQ7 streptococc
16	331	11.2	1142	2 Q93UZ9	Q93UZ9 bacillus sp

17	326.5	11.1	1287	2 Q9F930	Q9F930 streptococc
18	326	11.0	718	16 Q34587	Q34587 bacillus su
19	322.5	10.9	422	2 Q59243	Q59243 bacillus st
20	319.5	10.8	718	17 P95868	P95868 sulfolobus
21	311	10.5	718	2 Q69008	Q69008 thermus sp.
22	310.5	10.5	642	16 Q97Q88	Q97Q88 streptococc
23	306.5	10.4	666	16 Q84046	Q84046 chlamydia t
24	306	10.4	825	2 Q59319	Q59319 caidocellum
25	303.5	10.3	664	16 Q928F5	Q928F5 chlamydia p
26	302.5	10.2	1165	16 Q99XX8	Q99XX8 streptococc
27	302	10.2	783	10 Q04196	Q04196 arabidopsis
28	298.5	10.1	621	2 Q93HU3	Q93HU3 rhodothermu
29	298.5	10.1	716	17 Q973H3	Q973H3 sulfolobus
30	298	10.1	713	1 Q05152	Q05152 sulfolobus
31	296.5	10.0	818	10 Q41742	Q41742 zea mays (m
32	296	10.0	630	16 Q66936	Q66936 aquifex aeo
33	295	10.0	666	16 Q9PK26	Q9PK26 chlamydia m
34	295	10.0	789	10 Q22637	Q22637 zea mays (m
35	293.5	9.9	717	16 Q9K7U5	Q9K7U5 bacillus ha
36	292.5	9.9	660	1 Q9HHC8	Q9HHC8 thermococcu
37	292.5	9.9	741	3 Q9P5P3	Q9P5P3 neurospora
38	292	9.9	1072	16 Q9K6N1	Q9K6N1 bacillus ha
39	289	9.8	720	16 Q97FP9	Q97FP9 clostridium
40	289	9.8	741	2 Q59832	Q59832 streptomyce
41	288	9.7	741	2 Q9KY06	Q9KY06 streptomyce
42	286.5	9.7	666	16 Q9KNE8	Q9KNE8 vibrio chol
43	286.5	9.7	668	2 P71095	P71095 bacteroides
44	284	9.6	733	10 Q80403	Q80403 oryza sativ
45	284	9.6	1938	2 P70983	P70983 bacillus sp

ALIGNMENTS

RESULT 1

Q53641	PRELIMINARY;	PRT;	556 AA.
ID	Q53641; O08279; O08064;		
AC	Q53641; O08279; O08064;		
DT	01-NOV-1996 (TReMBLrel. 01, Created)		
DT	01-NOV-1996 (TReMBLrel. 01, Last sequence update)		
DT	01-DEC-2001 (TReMBLrel. 19, Last annotation update)		
DE	ALPHA-AMYLASE (FRAGMENT).		
GN	TREZ.		
OS	Sulfolobus acidocaldarius.		
OC	Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.		
OX	NCBI_TaxID=2285;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Kato M., Kettoku M., Miura Y., Komeda T., Konishi Y., Shindo K.,		
RA	Kobayashi K., Iwamatsu A.;		
RL	Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-ATCC33909;		
RX	MEDLINE=97135071; PubMed=8980629;		
RA	Maruta K., Mitsuzumi H., Nakada T., Kubota M., Chaen H., Fukuda S.,		
RA	Sugimoto T., Kurimoto M.;		
RT	"Cloning and sequencing of a cluster of genes encoding novel enzymes		
RT	of trehalose biosynthesis from thermophilic archaebacterium Sulfolobus		
RT	acidocaldarius.";		
RL	Biochim. Biophys. Acta 1291:177-181(1996).		
DR	EMBL; D64131; BAAL101.1; "		
DR	EMBL; D83245; BAAL1863.1; "		
DR	InterPro; IPR000461; Alpha-amylase.		
DR	InterPro; IPR004193; isoamylase_N.		
DR	InterPro; IPR003662; sub_transporter.		
DR	Pfam; PF00128; alpha-amylase; 1.		
DR	Pfam; PF02922; isoamylase N; 1.		
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.		
KW	Hydrolase.		
FT	NON_TER	556	556
SQ	SEQUENCE	556 AA;	64373 MW; 4215B45E6C8ED4E7 CRC64;

```
Query Match      100.0%; Score 2954; DB 1; Length 556;
Best Local Similarity 100.0%; Pred. No. 4.7e-189;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFSGGNTEKKNGIFKLWAPYVNSVKLSKLLIPMEKNDEGFEVEIDDEENLTYSYI 60
Db 1 MFSGGNTEKKNGIFKLWAPYVNSVKLSKLLIPMEKNDEGFEVEIDDEENLTYSYI 60

Qy 61 IEDKREIPDPASRYOPLGVHDKSOLIRTDYQILDGLGVKIEDLIYELHVGTFSEOGNFK 120
Db 61 IEDKREIPDPASRYOPLGVHDKSOLIRTDYQILDGLGVKIEDLIYELHVGTFSEOGNFK 120

Qy 121 GVIEKLDYLKDLGITGIELMPVAOPGPNRDWGDYGVFLYAVONTYGGPWEKAKLVNEAHK 180
Db 121 GVIEKLDYLKDLGITGIELMPVAOPGPNRDWGDYGVFLYAVONTYGGPWEKAKLVNEAHK 180

Qy 181 RGIIVILDVYVNHIGPEGNYLLGLGPFSDRYKTPWGLTFNFDGDCDQVRKFLENVEY 240
Db 181 RGIIVILDVYVNHIGPEGNYLLGLGPFSDRYKTPWGLTFNFDGDCDQVRKFLENVEY 240

Qy 241 WFKTFKIDGLRLDAVHAIFDPSPHILQEIATKAHQLGKFKFVIAESDLNDPKIVKDDCGYK 300
Db 241 WFKTFKIDGLRLDAVHAIFDPSPHILQEIATKAHQLGKFKFVIAESDLNDPKIVKDDCGYK 300

Qy 301 IDAQWVDDFHAVHAFITKEKDYIYQDFGRIDIEKTFKDVYVYDGKYSRYRGRTHGAPV 360
Db 301 IDAQWVDDFHAVHAFITKEKDYIYQDFGRIDIEKTFKDVYVYDGKYSRYRGRTHGAPV 360

Qy 361 GDLPPRFVFIQNHQDVGNRNGERLSILTDTKTYLMAATLYILSPYIPIFMGEYYE 420
Db 361 GDLPPRFVFIQNHQDVGNRNGERLSILTDTKTYLMAATLYILSPYIPIFMGEYYE 420

Qy 421 TNPEFFSDFSDPVLIKGVREGRLKENNQMDPOSEAFKSKLSWKIDEBVLDYKQLI 480
Db 421 TNPEFFSDFSDPVLIKGVREGRLKENNQMDPOSEAFKSKLSWKIDEBVLDYKQLI 480

Qy 481 NIKRYNNCKRKEVREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540
Db 481 NIKRYNNCKRKEVREGNCITLIMEKIGIIASFDDIVINSKITGNLLIGIGFPKKLKK 540

Qy 541 DELIKVNRGVGYOLE 556
Db 541 DELIKVNRGVGYOLE 556

RESULT 2
P95867 PRELIMINARY; PRT; 561 AA.
AC P95867;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE PRECURSOR (NALTO-OLIGOSYLITREHALOSE TREHALOHYDROLASE)
DE (TREZ) (EC 3.2.1.141).
GN TREZ.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweez M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Etraus G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charliebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RA "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
RL EMBL; Y08256; CNA6503.1; -.
```

```
DR EMBL; AE006815; AAK42272.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR00193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Hydrolase; Glycosidase; Complete proteome.
SQ SEQUENCE 561 AA; 64370 MW; B00E403020F6B242 CRC64;

Query Match      60.4%; Score 1784.5; DB 17; Length 561;
Best Local Similarity 59.2%; Pred. No. 4.9e-111;
Matches 332; Conservative 90; Mismatches 130; Indels 9; Gaps 6;

Qy 3 SFGNIEKNKNGIFKLWAPYVNSVKL-LSKLLIPMEKNDEGFEVEIDDEENLTYSYII 61
Db 2 TFGYKLEDDGVTFLWAPYQKVKILNRGIYEMERDDKGYFTITLNDVRVGRYKYL 61

Qy 62 EDKREIPDPASRYOPLGVHDKSOLIRTDYQILDGLGVKII--EDLIYELHVGTFSEOGNF 119
Db 62 DNSEVPDPASRYOPEGVHGYSIISPDFEWDNDNSVKVRKREDLVYIELHIGTFTSEGT 121

Qy 120 KGVIEKLDYLKDLGITGIELMPVAOPGPNRDWGDYGVFLYAVONTYGGPWEKAKLVNEAH 179
Db 122 EGVTRKUNYKELGVTAIEIMPIAQFFPKDWDGVTGLYAVQNSYGGPSGFRKLNVNEAH 181

Qy 180 KRGIAVILDVYVNHIGPEGNYLLGLGPFSDRYKTPWGLTFNFDGDCDQVRKFLENVE 239
Db 182 KGLAVILDVYVNHIGPEGNYVMKLGPFSEKYPKTPWGLTFNFDGDCDQVRKFLENVE 241

Qy 240 YWFTFKIDGLRLDAVHAIFDPSPHILQEIATKAHQLGKFKFVIAESDLNDPKIY--KDDC 297
Db 242 YWINEFVDDFRLDAVHAITDINSKPHILEDIADVVYHDKYIVIAESDLNDPVRVNPKEC 301

Qy 298 GYKIDAQWVDDFHAVHAFITKEKDYIYQDFGRIDIEKTFKDVYVYDGKYSRYRGRTHG 357
Db 302 GYNDADQWVDDFHAVHAFITGERQGYSDFGSGIDIVKSKYKIVIAESDLNDPVRVNPKEC 361

Qy 358 APVGDLPPRFVFIQNHQDVGNRNGERLSILTDTKTYLMAATLYILSPYIPIFMGEE 417
Db 362 KSVGDGLGDFVYVFIQNHQDVGNRNGERLKLVDKESYKIAAALYILSPYIPIFMGEE 421

Qy 418 YYETNPFFSDFSDPVLIKGVREGRLKENNQMDPOSEAFKSKLSWKIDEBVLDYK 477
Db 422 YGEENPFYIYDFSDPKLIQGVREGRRRNGQETDQSDCTFNDKSKLSWKINDILSYFK 481

Qy 478 QLINIRKRYN-NCKRKEVREGNCITLIMEKIGIIASFDDIVINSKITGNLLI--GTGF 534
Db 482 SLIKIRKEYGLACNRKLSVNGNWLIVKNGCLAVYVFSKSVIEMKYSGTLVLSNSSP 541

Qy 535 PKLKKDELKLVNRGVGYQL 555
Db 542 PSQITESK-VELDKGFALYKL 561

RESULT 3
Q55088 PRELIMINARY; PRT; 559 AA.
AC Q55088;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-AMYLASE.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KMI;
RA Kato M., Kettoku M., Miura Y., Komeda T., Konishi Y., Shindo K.,
RA Kobayashi K., Iwamatsu A.;
RA "The gene analysis of the new amylases from the hyper thermophilic
RT archaea Sulfolobus";
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
```

[illegible]

```

RESULT      5
Q973H4      PRELIMINARY;      PR2;      359 AA.
AC          Q973H4;
DT          01-DEC-2001 (TReMBLrel. 19, Created)
DT          01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT          01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE          PUTATIVE ALPHA-AMYLASE.
GN          ST0927.
OS          Sulfolobus tokodaii.
OC          Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX          NCBI_TaxId=111955;
RN          [1]
SEQUENCE FROM N.A.
RP          STRAIN=JCM 10545 / 7;
RC          PubMed=11572479;
RX          Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA          Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA          Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA          Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA          Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA          Oshima T., Kikuchi H.;

```

RT "Complete genome sequence of an aerobic thermophilic
RL Crenarchaeon, Sulfolobus tokodaii strain7.";
KW EMBL: AP000984; BAB565939.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 359 AA; 41600 MW; 4D9F429099E81665 CRC64;

Query Match 43.7%; Score 1290; DB 17; Length 359;
Best Local Similarity 66.0%; Pred. No. 2.6e-78;
Matches 235; Conservative 53; Mismatches 66; Indels 2; Gaps 2;
QY 201 LLGLGPFSDRYKTPWGLTNFDRGCDQVRKFTILENVEYWFKTKIDGLRLDAVHAIFD 260
DB 1 MMFLGPFYSQRYKTPWGLTNFDRGCDQVRKFTILENVEYWFKTKIDGLRLDAVHAIFD 60
QY 261 NSPKHILQETAETHAKHOLGKGVIAESDLNDPKIVKDCGKIDAOQWDDFHVAHFAFTKE 320
DB 61 SSPKHILODIAELSHSGKGVIAESDLNDPKIIDCKGKIDAOQWDDFHVAHFAFTKE 120
QY 321 KDYYQDFGRIEDIEKTFKDFVVDGYSRYRGRTHGAPVGDLPFRKFFVFIQNHDOGVN 380
DB 121 RNSYISDFGLDIDVAKDFVVDGYSRYRGRTHGAPVGNLACKFVVIQNHDOGVN 180
QY 381 RGNERISILTKTYTLYMAATLYTLSPYIPLFMGEYETNPFFSDFSDPVLKGV 440
DB 181 RGNDRSLTLVDKQSYMASALYLLSPFIPMFMEGEYETNPFFSDFSDPVLKGV 240
QY 441 EGRLEKNNQMDPOSEAFKSKLSKIDEDVDLYKQINIRKRY--NNCKRYKEVRREG 499
DB 241 EGRURDNGQSIDPOSDFAFLKSLKSWIDKEMLEYTLKVRKEFNCKSRRYEVDKTD 300
QY 500 NCITLIMEKITGASFDIDVINSKITGNLLIGIG--FPKLLKDKLIKVRNGVGYYQ 554
DB 301 NWLVIGLEKFLVAFSDTGLLRVNGELLESTDFPKRIVSSKEIKVKGVGYYK 356

RESULT 6

Q9ADI5 Q9ADI5 PRELIMINARY; PRT; 581 AA.
AC Q9ADI5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE PUTATIVE ALPHA AMYLASE.
GN SCBAC1A6.02C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP Streptomyces coelicolor.
RC SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cardeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=97000351; PubMed=8843436;
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmid and a detailed genetic and physical map for
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL589708; CAC33923.1; -;
DR InterPro: IPR000461; Alpha-amylose.
DR Pfam: PF00128; alpha-amylose; 1.
SQ SEQUENCE 581 AA; 63952 MW; 8816DFF3BC1A3408 CRC64;

Query Match 31.9%; Score 942.5; DB 2; Length 581;
Best Local Similarity 38.5%; Pred. No. 7.5e-55;
Matches 194; Conservative 93; Mismatches 174; Indels 43; Gaps 10;
QY 15 FKLWAPYVNSVKLKLKSLKLPMEKND--GFFEVEIDIEENLTYSYIIEDKREIPDPA 72
DB 3 FEVWAPQAGRVTLKRCDCATRALERDPERPGWCGEA--PARDGSRYGFAVDGPPVLPDPRS 61
QY 73 RYQPLGVHDKSOLIRTDYQILDGKV-----KTEDLIILYELHVGTSQEGNFKGVI 123
DB 62 RROPDPGDLG-----AVVDHGYAWRTPWAGRPLPGGVLYELHVGTTTPEGTLDAAA 114
QY 124 EKLDYKLDLGTITELMPVAQFPGNRDWDGDFVLYAVQNTYGGPMPWELAKLVNEAHRGI 183
DB 115 DRLEHLVRLGVTHVELMPLCFPGRCRHWGEGYSLWAVEHPYGGPEALKRFVDRHDLGL 174
QY 184 AVILDVYNNHIGPEGNYLLGLGPFSDRYKTPWGLTNFDRGCDQVRKFTILENVEYWF 243
DB 175 GVVLDDVNNHILGSPGNHLPAGPYFTDTHHTPMGCVAVNLDAPGSDVRAVLVDSALAWLR 234
QY 244 TFKIDGLRLDAVHAIFDPSKHILOETAETHAKHOLGK-----FVIAESDLNDPKIV--KD 295
DB 235 DYRLDGLRLDAVHALADTRACHLEQLSTAVDGLAADLERPLFLIAESDLNDPRIITPRA 294
QY 296 DCGYKIDAOQWDDFHVAHFAITKEDYDYODFGR--IEDIEKTFKDFVVDGYSRYRG 353
DB 295 EGGLVHQAONDDFHHAALHAALGEGSQGYADFARDPLAALAKTLIRGYFDGTYSYFRG 354
QY 354 RTHGAPV--GDLPPRFVFIQNHQDQVNGRNGERLSILTKTYTLYMAATLYTLSPYIPL 411
DB 355 RSHGRALDRKVAARHLTGYSQTHDQVGNRAQDRLASVSPGLAACATLTLTAPPTPM 414
QY 412 IFWGEYETNPFFSDFSDPVLKGVREGRLKE-----NNQMIDPOSEAFKSKL 464
DB 415 LFMGEWAAGTPWQFFTDHTDQLAQAVRQGRREFAAHGWREEDVDPDQDPATRESRCL 474
QY 465 SWKIDE-----EVLDDYKQINIR 483
DB 475 DWSEPEREPHARVLDWYRRLIALR 498

RESULT 7

Q911V1 Q911V1 PRELIMINARY; PRT; 583 AA.
AC Q911V1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE PROBABLE GLYCOSYL HYDROLASE.
GN PA2164.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL: AE004643; AAG05552.1; -;
DR InterPro: IPR002086; Aldehyde dehydr.
DR Pfam: PF00128; Alpha-amylose.
SQ SEQUENCE 583 AA; 63952 MW; 8816DFF3BC1A3408 CRC64;

RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL603646; CAC49676.1; -.
KW Transferase; Glycosyltransferase; Plasmid; Hypothetical protein;
KW Complete proteome.
SQ SEQUENCE 601 AA; 66954 MW; 024DA322A7B72C2E CRC64;

Query Match 26.8%; Score 792.5; DB 16; Length 601;
Best Local Similarity 36.9%; Pred. No. 8e-45;
Matches 192; Conservative 79; Mismatches 199; Indels 51; Gaps 15;

QY 3 SFGGN-IEKNKGIFKLPAPYNSVKLKLKSLIPMEKNDGFFVEIDDIENLTYSVII 61
DB 13 SWGANFIDSDTCRRLMAPDBREVDLVGGAVHKMSLDGGWFETTL-AAKAGERYCFRL 71
QY 62 EDKREIPDPASRYQPLGVHDKSOLI---RTDYLQILDGKVKIEDLIIVELHVGTFSEGN 118
DB 72 ADGTEVADPASSAQERASGTSIVVDAAYEWQTSRWGRPWEEAVISELHVGCFTPEGT 131
QY 119 PKGVIEKLDYLKDLGITGIELMPVAQPPGNRDWGYDGVFLYAVONTYGGPWELAKLVNEA 178
DB 132 FRAAIERLPHLAGAGITAIEIMPVAQPPGVGRWGYDGVFLYAPHNAYGKPDCLKALVDA 191
QY 179 HKRGIAVILDVYVNHIGPEGNYLLGLGP-YPEDRYKTPWGLTFNFDNRGCDQVRKFIEN 237
DB 192 HSLGLTVLLDVYVNHFGPEGNYLSRYASRFNKDRPTPWGASIAFEE---EAVRRYFIEN 248
QY 238 VEYWFKTFKIDGLRDVAHAFDINSFKHIL---OETAER-AHOLGKFVIAESDLNDPKI 292
DB 249 ALYWLGHFRDGLRLDATEQIORTKPHFLVALEHEVREAFARQIHLVLEDAHRRSL 308
QY 293 VKDCGKYI--DAQVDDFHAVHAFITKEDYIYQDF-----GRIEDIEKTFKDFVYD 345
DB 309 QRDASGARMFLFDAANDDLNHALHWATGETKGHYRLFADPEWCK---IRSALEAGFAPV 365
QY 346 GKYSRYGRTHGAPVD---LPPRKVFVFIONHDQVGNRNGERLSILTDTKTYLMAATL 402
DB 366 AKEDNF-----SPEGRARVPPOGRVNFLONHQDIGNRPAFGERLASLQEDSLRVLA 419
QY 403 YILSPYILPFMGEEYETNPFPPFSDPVLKVGREGRL-----KENNQ 450
DB 420 HMLTPQIPLLFMGEYGETQPFYFSDYQGEI-AAARLGRDRDAENFGGLPEGKTWDDL 478
QY 451 IDPOSEAFUKSLWSK-----IDDEVLYYKOLINTKRY 486
DB 479 PDPLDPDVFAGSKLRNRRATSPAGERHLAYVRDLAVIRQRH 519

RESULT 10
QY973H5 PRELIMINARY; PRT; 217 AA.
AC Q973H5;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE PUTATIVE ALPHA-AMYLASE.
GN ST0926.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Toshiyawa T., Tanaka T., Kudo H., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).

DR EMBL; AP000984; BAB65938.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 217 AA; 25181 MW; 4A1235EC5B491CEB CRC64;

Query Match 21.8%; Score 645; DB 17; Length 217;
Best Local Similarity 58.9%; Pred. No. 1.4e-35;
Matches 122; Conservative 37; Mismatches 44; Indels 4; Gaps 4;

QY 4 FGGN-IEKNKGIFKLPAPYNSVKLKLKSLIPMEKNDGFFVEIDDIENLTYSVII 62
DB 2 FGPSFLENEVEFILMAPYQKATRLNEEKYDMEKDEKGYFRITV-NAKEDKYSFII- 59
QY 63 DKREIPDPASRYQPLGVHDKSOLIIRTDYQILDG-KVKIEDLIIVELHVGTFSEGNFKG 121
DB 60 DKGEIPDPASRYQPDGVHGRSOLISLKYEWKSRVKKIPKEDLIIVELHIGTFQEGTFYS 119
QY 122 VIEKLDYLKDLGITGIELMPVAQPPGNRDWGYDGVFLYAVONTYGGPWELAKLVNEAHR 181
DB 120 AVEKLDYLVNLGVAIEIMPVHQFPKKDMGYDGVLYAVONTYGGPYAFMKFIDEAHR 179
QY 182 GIAVILDVYVNHIGPEGNYLLGLGPYF 208
DB 180 GLAVILDVYVNHVGPEGNYMMFLGPY 206

RESULT 11
Q44528 PRELIMINARY; PRT; 552 AA.
AC Q44528;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PUTATIVE ALPHA-GLUCANOTRANSFERASE /HYDROLASE.
GN LTI2.
OS Anabaena variabilis.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=1172;
RN [1]
RC STRAIN=M3;
RX MEDLINE=92119230; PubMed=1731972;
RA Sato N.;
RT "Cloning of a low-temperature-induced gene lti2 from the
cyanobacterium Anabaena variabilis M3 that is homologous to alpha-
amylases.";
RL Plant Mol. Biol. 18:165-170(1992).
DR EMBL; X59279; CAA41970.1; -.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Transferase; Hydrolase.
SQ SEQUENCE 552 AA; 63833 MW; 069886FAFD8FC411 CRC64;

Query Match 16.7%; Score 493; DB 2; Length 552;
Best Local Similarity 28.6%; Pred. No. 6.8e-25;
Matches 159; Conservative 76; Mismatches 142; Indels 178; Gaps 26;

QY 15 FKLWAPYNSVKLKLKSLIPMEKNDGFFVEIDDIENLTYSIIEDK 64
DB 7 FTLFAPYNGAALIASFSDMQEIPMKKGGDGYFRTTVELEDGYQYKFRVQTRSWFFED 66
QY 65 R--EIPDPASR--YQPLGVHDKSOLIIRTDYQILDGKVKIED-----LIIVELHVG 112
DB 67 QWVDVTPYATIDESSGKDNSTARKGEIKIVDTYVWQHDDKPLPADHELVIYELHVG 126
QY 113 FS-----QEGNFKGVIEKLDYKDLGITGIELMPVAQPPGNRDWGYDGVFLYAVONT 165
DB 127 FSGGEDDPYARGKYKHVIEKLDYLCELGINAIELLPVKEYPGDSWGNRPYFATESSY 186
QY 166 GGPWELAKLVNEAHRKRGIAVILDVYVNH-----GPEGNYLLGLG 205


```
Db 187 GSTADLKKLVDECHQIRIIMDGIYNHSEASSPLTQIDHDYVWHHEPRDPDNN-----240
QY 206 PYFSRDYKTPGWTGFNFDD--DRGCD--QVRKFEILENVYWFYKFKIDGLRLDAVHAIFDN 261
Db 241 -----WGPEFNTEHYDENLEYTPARKFIGDVIYVWGEYHLDGIRYDAARQIANY 290
QY 262 SPKH-ILQETAEKAHOLGKEVIAE-----SDLNDPKIVKDCGKYKIDAOQVDDFHHA 312
Db 291 DEWHIAQEAKKTAGAKPFYVNAEHIPETTSIINLGP-----MDGCVHDSFYHT 340
QY 313 VHAFITKEKDYIQDFGRIEDIEKTKDFVYDGKYSRYGRTHGAPVGLDPPRKFFVFI 372
Db 341 IKAHIC-----GDTFDLE-NLKDVI--DPKROGFLGATN-----VVNYL 376
QY 373 QNHDD-----QVNGRNGERLSILTDK-----TTLMAATLYILSPYIPIEWMGEY 418
Db 377 TNDHHDHIMVGLNR-----EFHDEAFRAKLGTAILMTAV-----GVPLIMWGEF 424
QY 419 YETNPFFFSDFSDPVLKGVREGRLKENNMIDPOSEEAFKLSLW-----KIDEVL 473
Db 425 GEYKP-----KQDQ-----SKIDWTLGLGDLNRLSLF 451
QY 474 DYYKQLINIRKRYNN 488
Db 452 DYHKGLIGLRK--NN 464

RESULT 12
Q93035 PRELIMINARY; PRT; 440 AA.
AC Q93035;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE BRANCHING ENZYME GLGB (FRAGMENT).
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RA Ueki T., Inouye S.;
RT Identification of a new His-Asp phosphorelay signal transduction
RT system which regulates expression of a heat shock gene, lond, of
RT Myxococcus xanthus.
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF285783; AAK83002.1;
FT NON_TER 440
SQ SEQUENCE 440 AA; 48964 MW; 8D492EA5A492017 CRC64;

Query Match 14.2%; Score 420; DB 2; Length 440;
Best Local Similarity 32.0%; Pred. No. 3.7e-20;
Matches 133; Conservative 52; Mismatches 129; Indels 102; Gaps 17;

QY 52 EENLTYIIEEDKREIPDPASRYQPLGVDKSKLSQITDYQILDGKVKIEDLIYELHVG 111
Db 80 QENSTGSSIIYDHGEYWNAAQYSSPG-----FNEMLIYEMHVG 118
QY 112 TFSQE-----GNPKGVIEKLDYLDKLGITGIELMPVAQPPGNGRDWGYGVFLYAVONTYG 166
Db 119 TPHDSPGPGNGNSAIARLDHVRDLGANNIKVMPAYEFAGDFSWGYNAAFPPAPESAYG 178
QY 167 GPWELAKLVNEAHKRGIAVILVVYVNHGP-----EGNYLLGLGYPFSD--RYKTPW 216
Db 179 HPNDMKRFYDEAHMRGIGVIFDVVYVNHGYSDLPMCMFCGDCILGSGGEYFYNDWRKSTPW 238
QY 217 GLTFNFDRCQDQVRKFEILENVYWFYKFKIDGLRLDAVH-----AIFD--NSP 263
Db 239 GDT--RPDGRPRVIRAYIRDSMNMLTSPFGGLRWDAIKYMTQNGSTDTAIPDAWRVF 296
QY 264 KHILQETAEKAHOLGKFEVIAESDLNDPKIVKDDC-----GYKIDAOQVDDFHHAHVAFI 317
```

```
Db 297 RSINREI--NATPWKISIAE-DFGGDFITNDATSDTSGAGFDSQWGGDFVHAIRAAV 353
QY 318 TREKDYIYQDFGRIEDIEKTKDFVYDGKYSRYGRTHGAPVGLDPPRFVYFIQNHQD 377
Db 354 IAS-----NDSGR--DMNS-----VRNATQYSGR-HTAR-----VIYSESHDE 390
QY 378 V-----GNGRNGERLSILTDKTYLMAATLYILSPYIPIEWMGEY 420
Db 391 VANGKARVPEEINPGNAGS-----WAAKARSLAAGVETSPGIPMIFOGQEFL 440

RESULT 13
Q59242 PRELIMINARY; PRT; 652 AA.
AC Q59242; O08511; O08495;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE 1,4-ALPHA-GLUCAN BRANCHING ENZYME (EC 2.4.1.18)
DE (GLYCOGEN BRANCHING ENZYME) (1,4-ALPHA-GLUCAN BRANCHING ENZYME)
DE (AMYLO-(1,4 TO 1,6)TRANSGLUCOSIDASE)
DE (AMYLO-(1,4-1,6)-TRANSGLYCOSYLASE).
GN GLGB.
OS Bacillus steatothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RX MEDLINE=95031021; PubMed=7944355;
RA Takata H., Takaha T., Kuriki T., Okada S., Takagi M., Imanaka T.;
RT Properties and active center of the thermostable branching enzyme
RT from Bacillus steatothermophilus.
RL Appl. Environ. Microbiol. 60:3096-3104(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RA Takata H.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=TRBE14;
RA Takata H., Takaha T., Okada S., Takagi M., Imanaka T.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: FORMATION OF 1,6-GLUCOSIDIC LINKAGES OF
CC GLYCOGEN.
CC -1- PATHWAY: THIRD STEP IN GLYCOGEN BIOSYNTHESIS.
CC -1- MISCELLANEOUS: OPTIMAL ACTIVITY AT APPROXIMATELY 39 DEGREES
CC CELSIUS.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
DR EMBL; D87026; BAA19588.1;
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; isoamylase_N.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02922; isoamylase_N; 1.
KW Glycogen biosynthesis; transferase; Glycosyltransferase.
FT ACT_SITE 308 308 BY SIMILARITY.
FT ACT_SITE 351 351 BY SIMILARITY.
FT ACT_SITE 419 419 BY SIMILARITY.
SQ SEQUENCE 652 AA; 76793 MW; 4591EB414A0E3FEF CRC64;

Query Match 12.2%; Score 360; DB 2; Length 652;
Best Local Similarity 25.3%; Pred. No. 6.3e-16;
Matches 142; Conservative 72; Mismatches 186; Indels 162; Gaps 25;

QY 4 FGNIENKNG-----IFKLWAPYVNSVKL-----KLSKLIPEKNDGFEVE 47
Db 26 FGAHVINEGGKVGTRFCVWAPHAREVRLVGSFNDWGDGDFRLEK-----VNDEGVMTIV 79
```

```
Qy 48 IDDIENLTYSY-IIDKREI---PDPASRYQPLGVHDKSOLIRTD---YQILDIG----- 96
Db 80 VPENLEGLHYEIVTPDQVLFRAFPYAFYSELPHAS---IAYDLKAGOWNDSWKRK 137
Qy 97 ---KVKIEDLIIYELHVGTF-SQEGNFKGVIEKLD-----YKDLGITGIELMPVAQPPG 147
Db 138 KRRRIYDQPMVYIELHFGSKKDGREYTYREMADELISYVDHGFTHEILLPLVEHPL 197
Qy 148 NRDMGYDGVFLYAVONTYGGWELAKLNEAHKRGIAVILDVYVNIH--GPEGNYLLGLG 205
Db 198 DRSMGYQGTGYAYTSYGTGTHDPMYFYDRCHQAGIGVMDWPGCHFKDAHGLYMDGA 257
Qy 206 PYF-----SDRYKTPWGLTWNFDRCQDVRRKFIENVEYWFKTKIDGLRLDVA----- 255
Db 258 PTYEYANEKDRENYWG-TANF-DLQKPEVRSFLISNALFWLEYHYHIDGFRVDAVANMLY 315
Qy 256 ---HAIFDNS-PKHILQIEAEK--AHQLGKFVTAESDLNDPKIVKD-----DCGYKID 302
Db 316 WPNDRLYENPYAVEFLSKLNEAVFAYDPNALMTAEDSTDWPKVTAPTTEGGGLGFNYKWN 375
Qy 303 AQWVDDF-----HHAHVAFITTEKDYQYQDFGRIEDIEKTFKDVYVDGKYSRYR 352
Db 376 MGWMDMLKYMETPPYERRHVHNVQTFSLAYS-----ENFILPFSHDEVVHGKKS--- 427
Qy 353 GRTHGAPVDLPPRKFFVFIQNHQVGNRNG-----ERLSILTDKTYLMAATLYILSPYI 409
Db 428 -----LLNKMPGSVEERKFAQLRLLYGYMA-----HPGK 456
Qy 410 PLIFMGEYETNPPFFSDFSDPVLKGVREGLRKNNQMDPQSEAFLLKSLW--- 466
Db 457 KLLFMGNEF-----AQDFDEWKEFEDELDDWLF 482
Qy 467 --KIDEEVLDYKQLINIRKRY 486
Db 483 DFLHRKNDYMKELIACYKRY 504
RESULT 14
P73608 ID P73608 PRELIMINARY; PRT; 707 AA.
AC P73608:
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GLYCOGEN OPERON PROTEIN GLGX.
GN GLGX OR SLR1857.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Suglura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90908; BAA17652.1; -.
DR HSP; P10342; 1BF2.
DR InterPro; IPR000461; Alpha_amylase.
DR InterPro; IPR004193; Isoamylase-N.
DR pfam; PF00128; alpha-amylase_1.
DR pfam; PF02922; Isoamylase_N; 1.
KW Complete proteome.
SQ SEQUENCE 707 AA; 79895 MW; CFA27B3C86318DB1 CRC64;
```

Query Match 11.38; Score 333.5; DB 16; Length 707;
Best Local Similarity 23.18; Pred. No. 4.1e-14;

```
Matches 146; Conservative 83; Mismatches 197; Indels 207; Gaps 29;
Qy 2 FSGFIEKKNKGIKFLWAPYVNSVKLSKLLIPMEKNDEGFEVE----- 47
Db 23 FPGATIVPGGVNSIYSSHSTACTLVLFKRAP-----QPFVEIPPESEFRIGNVYCMV 77
Qy 48 IDDIENLTYSY-IIDKREI---PDPASRYQPLGVHDKSOLIRTDYQILDIG----- 97
Db 78 VFDLDFENLEYGYRMEG-----PNNFQOQHWFDPKSVLLDPYAKVVSGRDVGWGTQNW 130
Qy 98 -----VKIEDLIIYELHVGTFSQE-----GNFKQVI 123
Db 131 DDIYQHRGRUSFDDFWDENSDPLDVPLEDMVYIEMHVHVGFTKDPSSGYSKENVHRTFAGIL 190
Qy 124 EKLDYLDLGTGTIELMPVAQF-----PGNRD-----WGYDGVFLYAVONTYGGPW 169
Db 191 SKIPLYQLGLVNTIELMPIFEDEFEEHRSYHPETGEFLVNWYGVSTVNFPAKAGYATG 250
Qy 170 -----ELAKLVNEAHKRGIAVILDVYVNIHIGPEGNYLLGLGYPFSDR-----YKT 214
Db 251 KFGWQIDELKNLVKELHKGISVILDVVFNHTA-EGN---ERGPTISFRGLDNKTYMLT 306
Qy 215 PWGLTFNFDNRG---CDQ---VRKFIENVEYWFKTKIDGLRLDVAHAIQDNPKHILQ 268
Db 307 PEGYFNFSGTGNLNCNNPIVRGMVLDCLRYTAEFHIDGFRD-LASILGRDP----- 360
Qy 269 EIAEKAHOLGKFVTAESDLNDP---KIVKD---DCG--YKID-----AQMVDFFH 311
Db 361 ---WGYPLANPPLLETLAFDPILARSKLIAEAWDAGLYQVGSFSGRWAENNGKYRD 416
Qy 312 AVHAFITTEKDY-----YYDFGRIEDIEKTF-----KDFVYVDGKYS 349
Db 417 TVRKFIKGDAGVICEMAORLQSGPDLYOGAGRPSTSFNFVTAHGFADLADLVAYNGKH 476
Qy 350 RYRGTHGAPVDLPPRKFFVFIQNHQVGNRNGERLSILTDKTYLM-AATLYILSP 407
Db 477 YANGENGNDGAND-----NYSWNCVGEPTDNPDLRLRARQMRNAIALLVSQ 525
Qy 408 YIPLIFMGE-----YYETNPPFFSDFSDPVLKGVREGLRKNNQMDIDPOSE 456
Db 526 GVPMLMGDEMGKTDGNNNTYCHDSPFNWLN-----WHLEQN-----QAW 567
Qy 457 EAFLLKSLKSKIDEVLDYKQLINIRKRYNNC 489
Db 568 FRFVKHCIAFRLAHPVL-----RNSEHFQNC 593
RESULT 15
Q97S07 ID Q97S07 PRELIMINARY; PRT; 1280 AA.
AC Q97S07:
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALKALINE AMYLOPULLANASE, PUTATIVE.
GN SP0268.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=TiGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Maynam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple L., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
```


